Dual-energy computed tomography for proton radiotherapy

Esther Bär

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

at the

University College London.

Department of Medical Physics and Biomedical Engineering

Proton and Advanced Radiotherapy Group

University College London

October 2, 2018

Supervisors: Hugo Bouchard – Department of Physics, Université de Montréal
Gary Royle – Proton and Advanced Radiotherapy Group, UCL

I, Esther Bär, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

Proton therapy is an advanced form of cancer treatment. To use this precise technique to its full potential, a better understanding of the uncertainties involved in the treatment process and their mitigation is necessary. This thesis aims at the investigation, understanding, and improvement of three sources of range uncertainties: (i) CT imaging and conversion to tissue, (ii) mean excitation energies and (iii) lateral inhomogeneities in the irradiated object.

To predict the range of the proton beam within the patient, knowledge of the tissues in the beam's path is required. Clinically, the required tissue characteristics are estimated using a single-energy CT (SECT) scan of the patient. In this work, the potential of dual-energy CT (DECT) to improve the estimation of tissue characteristics for proton therapy planning and the related uncertainty is shown in two steps. In a first study, several DECT approaches are compared in a theoretical study and their performance in realistic situations is evaluated. In a second study, DECT-predicted tissue characteristics are validated in an experimental setup using animal tissues.

To understand the influence of the mean excitation energy (I-value) on the range uncertainty, elemental I-values for particle therapy planning are revisited and an uncertainty budget is established. This enables the estimation of stopping power and range uncertainties arising from I-values.

The largest source of range uncertainties in particle therapy planning arises from lateral inhomogeneities in the irradiated body, introducing range degradation. Conventional dose calculation algorithms use ray tracing to cal-

culate the beam range, leading to severe range uncertainties. Monte Carlo (MC) was demonstrated to reduce these uncertainties by accurately simulating particle transport in the patient geometry. In this work, the uncertainties arising from lateral inhomogeneities are investigated with both, ray tracing and MC techniques.

Acknowledgements

This thesis does not only present my work but involved the contributions of many collaborators. I was extremely lucky to spend time at different prestigious institutions across the world and work with experts in their field, who kindly shared their knowledge and provided me with what I needed to complete this PhD. My experience throughout this journey has been nothing but amazing.

First and foremost I would like to express gratitude to my supervisors, Hugo Bouchard and Gary Royle. Hugo Bouchard has provided the most valuable guidance and mentoring. I am grateful for all the knowledge he shared with me, for the hours spent on discussion, for all the patience, motivation, enthusiasm, encouragement and trust. Thank you for believing in me, even when I couldn't. I will always see you as a supervisor.

I would like to express my sincerest gratitude to Gary Royle for making this PhD a great experience. Gary made it possible for me to conduct experiments and gain valuable experience abroad. Without his constant support, guidance, and trust, this work would not have been possible.

I owe a very special thank you to my colleague, close collaborator, and friend Arthur Lalonde. Thank you for the countless hours spent on discussing details of our projects and for being involved with my work. I could not have wished for a better colleague to share projects and ideas with. But most of all, thank you for your friendship, which has been invaluable during the past

three years. Thank you for the hours spent at pubs, watching hockey games, running or ice skating, and for showing me the best sides of Montreal.

I would like to thank the Medical Radiation Physics and Medical Radiation Science teams at NPL for hosting me during my first year. Especially Rebecca Nutbrown, who made it possible for me to spend time at NPL. I am extremely grateful for this opportunity, and for the resources and funding provided by NPL. Thanks to all the group members for the good times, for many chats during coffee or pub nights.

I would like to express my appreciation to the people I worked with at MGH. To Hsiao-Ming Lu for collaborating with me and mentoring the experimental part of the work. To my colleagues Rongxiao Zhang, KyungWook Jee, Kai Yang and Gregory Sharp for their involvement in the experiments. I am grateful to Marta Dias, Lennart Volz and Hakan Östen, who grew close to me during my time in Boston and made the days at the office so much more enjoyable. My most heartfelt appreciation goes to Charles-Antoine Collins-Fekete, who has been my confidant from the day we met. Thank you for being there for me, even during the worst of my times. You have encouraged and reassured me a billion times at least, and you always know how to cheer me up. I can't express how grateful I am.

I am deeply grateful to my co-workers from the Medical Physics group at the CHUM in Montreal. I was lucky to share the office with amazing people who made my Canadian experience incredibly enjoyable. I'd like to start with Mikaël Simard, who was a very close friend from the moment we met. Thank you for all the talks, for brightening up my days, for running, for the nights at the bars and pubs. I would also like to thank Andréanne Lapointe, Charlotte Remy, Jasmine Duchaine and Yunuen Cervantes for the good times, and for including me in the group. I would also like to mention the friends I made at

the German Stammtisch in Montreal, who I have met almost weekly. Thank you for bringing a piece of Germany to Canada. I have very fond memories of my time in Montreal, and I owe them to all of you guys.

A big thank you to everyone in the Department of Medical Physics and Biomedical Engineering at UCL. Thank you to the #protongirls Reem Ahmad, Catarina Veiga, Ana-Monica Lourenço, Megan Wilson, Stacey McGowan and Savanna Chung for always being there for chats, drinks or coffee breaks. To Emma Biondetti, Paul Burke, Peter Modregger, Michael McManus, and Chi Shuo Chuang, for your friendship and support. I am feeling very lucky to have met each of you.

A very emotional thanks to my long-term friends Vânia Batista, Nico Schubach, Peter Hintz and Laura Schlinke. Even though we don't see each other often, knowing that you are there for me when needed has given me great support. A big thank you to Thomas Klinge, my closest friend. Knowing that I can always rely on you has been invaluable in the past two years. I would also like to express my appreciation to Christian Maaß. Thank you for being there for me throughout our good times.

I want to express the most special thanks to my family, especially to my parents Dirk and Sandra Bär and my siblings Elisa and Jonathan Bär, for their unconditional love and support.

Contents

List of Figures

List of Tables

Acronyms

AAPM	American Association of Physicists in Medicine
CDF	Cumulative density function
CSDA	Continuously slowing down approximation
CT	Computed tomography
CTV	Clinical target volume
DECT	Dual-energy computed tomography
DEI	Dual-energy index
EAN	Effective atomic number
ED	Electron density
ESTRO	European SocieTy for Radiotherapy and Oncology
ETD	Eigentissue decomposition
GTV	Gross tumor volume
HU	Hounsfield units
ICRU	International Commission on Radiation Units and Measurements
LWF	Local weighting function
MC	Monte Carlo
MECT	Multi-energy computed tomography
MGH	Massachusetts General Hospital
NPL	National Physical Laboratory

PDF	Probability density function
PTCOG	Particle Therapy Co-Operative Group
PTV	Planning target volume
PPRIG	Proton Physics Research Implementation Group
RSP	Relative stopping power
SECT	Single-energy computed tomography
SPR	Stopping power ratio
TOPAS	TOol for PArticle Simulation
WEPL	Water equivalent path length
WER	Water equivalent range

Publications

The work presented in this thesis has resulted in the following outputs.

Peer-reviewed journal papers:

- **Bär, E.**, Lalonde, A., Royle, G., Lu, H.-M. and Bouchard, H. (2017). The potential of dual-energy CT to reduce proton beam range uncertainties. *Medical Physics* 44: 2332-2344. doi:10.1002/mp.12215.

- Zang, R., **Bär, E.**, Jee, K. W., Sharp, G. C., Flanz, J., Lu, H.-M. (2017) Investigation of real tissue water equivalent path lengths using an efficient dose extinction method. *Phys. Med. Biol.* 62: 5640-5651. doi: 10.1088/1361-6560/aa782c.

- Lalonde, A., **Bär, E.** and Bouchard, H. (2017) A Bayesian approach to solve proton stopping powers from noisy multienergy CT data. *Medical Physics* 44: 5293-5302. doi: 10.1002/mp.12489.

- Collins-Fekete, C.-A., **Bär, E.**, Volz, L., Bouchard, H., Beaulieu, L. Seco, J. (2017) Extension of the FermiEyges most-likely path in heterogeneous medium with prior knowledge information. *Phys. Med. Biol.* 62: 9207-9219. doi: 10.1088/1361-6560/aa955d.

- **Bär, E.**, Lalonde, A., Zhang, R., Jee, K. W., Yang, K. , Sharp, G. C., Liu, B., Royle, G., Bouchard, H., Lu, H. M. (2018) Experimental validation of two dual-energy CT methods for proton therapy using heterogeneous tissue samples. *Medical Physics* 45: 48-59. doi: 10.1002/mp.12666.

- **Bär, E.**, Andreo, P., Lalonde, A., Royle, G., Bouchard, H. (2018) Optimized I-values for the use with the Bragg additivity rule and their impact on proton stopping power and range uncertainty. *Phys. Med. Biol.* 63: 165007-165024. doi: 10.1088/1361-6560/aad312.

- Lalonde, A., Simard, M., Remy, C., **Bär, E.**, Bouchard, H. (2018) The impact of multi-energy CT on proton pencil beam range uncertainties: a Monte Carlo study. *Phys. Med. Biol.*, in press.

Oral presentations:

- **Bär, E.**, Lalonde, A., Zhang, R., Jee, K. W., Yang, K. , Sharp, G. C., Liu, B., Royle, G., Bouchard, H., Lu, H. M. (2016) Experimental validation of the DECT stoichiometric calibration for proton treatment planning using real tissue samples. Boston, *AAPM New England Chapter, Peter Neurath Young Investigator Symposium.*

- **Bär, E.**, Lalonde, A., Zhang, R., Jee, K. W., Yang, K. , Sharp, G. C., Liu, B., Royle, G., Bouchard, H., Lu, H. M. (2016) The Impact of Using Dual-Energy CT for Determining Proton Stopping Powers: Comparison Between Theory and Experiments. Washington D.C., *58th Annual Meeting of the AAPM. Medical Physics* 43 (6Part35): 3756-3756. doi: 10.1118/1.4957542.

- **Bär, E.** Lalonde, A., Zhang, R., Jee, K. W., Yang, K. , Sharp, G. C., Liu, B., Royle, G., Bouchard, H., Lu, H. M. (2016). The Impact of Using Dual-Energy CT for Determining Proton Stopping Powers of Real Tissues. London, *3rd PPRIG workshop.*

- Lalonde, A., **Bär, E.** and Bouchard, H. (2017) A Bayesian method to derive proton stopping powers of human tissues from multienergy CT data. Denver, *59th Annual Meeting of the AAPM. Medical Physics* 44 (6): 3259.

- **Bär, E.**, Andreo, P., Lalonde, A., Royle, G., Bouchard, H., (2017) A new method to improve the accuracy of the Bragg additivity rule for calculating proton stopping powers. London, *4th PPRIG Workshop.*

- **Bär, E.**, Andreo, P., Lalonde, A., Royle, G., Bouchard, H., (2018) A novel method to estimate mean excitation energies and their uncertainties for particle therapy. Barcelona, *ESTRO 37. Radiother. Oncol.* 127: S42-S43.

- **Bär, E.**, Lalonde, A., Royle, G., Bouchard, H., and Lu, H.-M. (2018). Dual-energy CT to estimate proton stopping powers: A validation using real tissues. London, *ENLIGHT Annual Meeting.*

Poster presentations:

- **Bär, E.**, Lalonde, A., Royle, G., Lu, H.-M. and Bouchard, H. (2016). On the Performance of Four Dual Energy CT Formalisms for Extracting Proton Stopping Powers. Washington D.C., *58th Annual Meeting of the AAPM. Medical physics* 43 (6Part11): 3453-3453. doi: 10.1118/1.4956103.

- Zang, R., **Bär, E.**, Jee, K. W., Sharp, G. C., Flanz, J., Lu, H.-M. (2016) Efficient dose extinction method for water equivalent path length (WEPL) of real tissue samples for validation of CT HU to stopping power conversion. Washington D.C., *58th Annual Meeting of the AAPM. Medical Physics* 43(6Part11): 3452-3452. doi: 10.1118/1.4956101.

- **Bär, E.**, Lalonde, A., Royle, G., and Bouchard, H. (2016). On the performance of dual-energy CT for determining proton stopping powers. London, *3rd PPRIG workshop.*

- **Bär, E.**, Lalonde, A., Royle, G., Lu, H. M., Bouchard, H. (2017) On the performance of dual-energy CT for determining stopping power ratios for proton radiotherapy. Montreal, *9th QBIN Scientific Day.*

- **Bär, E.**, Lalonde, A., Royle, G., Bouchard, H. (2017) On the potential of dual-energy CT to predict nuclear interaction cross sections and radiation length of therapeutic proton beams. Denver, *59th Annual Meeting of the AAPM. Medical Physics* 44 (6): 2812-2813.

- Wang, H. Wu, J., Zang, R., **Bär, E.**, Geng, C., Jee, K. W., Sharp, G. C., Paganetti, H., Tang, J., Lu, H.-M. (2017) Validation of Monte-Carlo proton dose calculation for real tissue samples. Denver, *59th Annual Meeting of the AAPM. Medical Physics* 44(6): 28532854.

- Collins-Fekete, C.-A., **Bär, E.**, Volz, L., Bouchard, H., Beaulieu, L. Seco, J. (2017) Extension of the FermiEyges most-likely path in heterogeneous medium with prior knowledge information. Denver, *59th Annual Meeting of the AAPM. Medical Physics* 44 (6): 2770.

- Lalonde, A., Remy, C., **Bär, E.** and Bouchard, H. (2018) Monte Carlo validation of a new dual-energy CT method for proton therapy in a patient-like geometry. Barcelona, *ESTRO 37. Radiother. Oncol.* 127: S535-S536.

- **Bär, E.**, Lalonde, A., Royle, G., Bouchard, H., and Lu, H.-M. (2018). Dual-energy CT to estimate proton stopping powers: A validation using real tissues. London, *ENLIGHT Annual Meeting.*

- **Bär, E.**, Lalonde, A., Royle, G., Lu, H. M., Bouchard, H. (2018) Experimental quantification of range uncertainties arising from lateral inhomogeneities. Nashville, *60th Annual Meeting of the AAPM, presentation as electronic poster.*

Prizes

The work presented in this thesis has received the following prizes.

- 58th Annual Meeting of the American Association of Medical Physics (AAPM): The conference abstract *Impact of using dual-energy CT for determining proton stopping powers of real tissues* was selected as one of the best research works in the field of imaging for particle therapy and presented in the **Science Council Session** in August 2016.

- The paper entitled *Experimental validation of two dual-energy CT methods for proton therapy using heterogeneous tissue samples* was selected **Editor's choice** in *Medical Physics*.

- UCL Department of Medical Physics and Biomedical Engineering PhD Showcase: Winner of the **Presentation Style Award** in June 2018.

- ENLIGHT Annual Meeting 2018: The poster entitled *Dual-energy CT to estimate proton stopping powers: A validation using real tissues* received the **Prize for winning poster** and was selected for oral presentation in June 2018.

Chapter 1

Introduction

1.1 Radiotherapy with protons

Proton radiotherapy is a form of cancer treatment utilizing a beam of accelerated protons directed at the tumor. The main advantage of radiotherapy with protons over radiotherapy with photons lies in their favorable dose deposition pattern. Protons continuously slow down in many interactions with the electrons and nuclei of the target material, depositing dose along their paths. Due to the nature of the interactions, the energy loss and thus the energy deposited locally in the target material increases with decreasing proton velocity, leading to the formation of a well-defined dose peak, the so-called Bragg peak. A fraction of the protons do not contribute to the Bragg peak due to large angular scattering or interactions with the target nucleus. The Bragg peak is followed by a steep dose fall off, with very low dose behind it arising from neutrons and gamma radiation produced during nuclear interactions, giving the proton beam a finite range within the irradiated material. This finite range, together with the low entrance dose, results in a lower integral dose compared to photon radiotherapy for a given target dose, and to a better sparing of healthy tissues distal to the target volume.

In addition to the favorable depth dose deposition, another advantage of proton therapy over photon therapy is the sharp lateral dose profile. A sharp

lateral penumbra is important to spare organs next to the target volume. The achievable penumbra depends on several factors, such as beam delivery system and depth in the patient. Proton penumbras at shallow depths are steeper than photon penumbras. However, while photon penumbras show a moderate increase with depth, proton penumbras widen more drastically. For example, the penumbra of a proton beam of a range of 28 cm in water is larger than a 15 MV photon beam [1].

In recent years, proton therapy has established itself as a preferred treatment method for various indications. Those indications include, among others, tumors of the central nervous system, head and neck cancers, hepatocellular cancers, ocular tumors, base of skull tumors, and primary tumors in children [2].

As of 2016, a total of approximately 150,000 patients were treated with proton therapy (according to the Particle Therapy Co-Operative Group (PTCOG) Patient statistics per end of 2016) [3], with the number of treatment centers steadily growing. Currently, 79 particle treatment centers are operational, with another 46 under construction and 22 in the planning stage (as of March 2018) [4].

1.1.1 Proton interactions with matter

In the following, the physical interaction processes of a beam of accelerated protons with matter are described. In contrast to photons, protons and ions are not attenuated by the atoms of the traversed material but slow down in many interactions. The accelerated particle is referred to as the projectile, while the traversed material is the target. The most frequent interactions are collisions with shell electrons of the target material. This interaction leads to the production of secondary electrons and ionization of the target atom, while the projectile experiences a small amount of energy loss.

Another frequent interaction is scattering of the projectile off the electric field of a target nucleus. The positively charged ion and the atomic nucleus repel each other and the projectile gets deflected from its path. The scattering angle

of a single scattering event is usually small, as described by the Rutherford differential scattering cross section [5]. In thick absorbers, projectiles experience multiple Coulomb scattering (MCS) events, leading to a nearly Gaussian angular distribution behind the scatterer. In case of a pencil beam of particles, MCS will lead to a widening of the beam. The scattering probability and the angular distribution of a pencil beam behind a thick absorber are described by Moliére's theory [6].

A third group of interactions are collisions with the atomic nucleus. These nuclear interactions can be either elastic, nonelastic or inelastic. In an elastic nuclear interaction, the incident projectile scatters off the target while the state of target nucleus and projectile stay unchanged. In some cases, the collision of projectile and target leads to excitation of the target nucleus, referred to as inelastic nuclear interaction. In other cases, the collision leads to the fragmentation of the target nucleus, referred to as nonelastic nuclear interaction. Particles that underwent inelastic or nonelastic nuclear interactions are called secondary particles.

1.1.2 Energy loss of protons

Protons traversing a slab of material loose their energy in many interactions with the atoms of the target material, until coming to rest. The mean energy loss $\mathrm{d}E$ of a beam of particles traversing through a slab of material $\mathrm{d}x$ is defined as

$$S = \frac{\mathrm{d}E}{\mathrm{d}x}, \tag{1.1}$$

where S is the linear stopping power of the material. The stopping power consists of several components accounting for electronic (el), radiative (rad) and nuclear (nuc) effects

$$S = \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\mathrm{el}} + \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\mathrm{rad}} + \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\mathrm{nuc}}. \tag{1.2}$$

For protons at therapeutic energies, the dominating contribution to the energy loss is the electronic stopping power, also called collision stopping power

$\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{el}$. The radiative stopping power accounts for radiative losses from Bremsstrahlung and is irrelevant for protons at therapeutic energies. The nuclear stopping power describes the energy loss from elastic Coulomb scattering and is only relevant at low proton energies. Please note that energy loss due to nuclear nonelastic interactions is usually not included in S.

The energy loss of particles was formulated by Bethe (1930) [7], with several corrections added retrospectively [8] and is described as

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = k\frac{z^2}{\beta^2}N_\mathrm{e}\left(\ln\frac{2m_\mathrm{e}c^2\beta^2}{I\left(1-\beta^2\right)} - \beta^2 - \frac{C}{Z} - \frac{\delta}{2}\right). \tag{1.3}$$

In the above equation, k is a constant, z the projectile charge ($z = 1$ for protons), β is the velocity of the projectile relative to the speed of light c, N_e is the electron density of the target medium, m_e is the electron rest mass, I is the mean excitation energy of the medium, C is the shell correction constant, Z the atomic number of the target material, and δ the density correction constant.

The most important property in the Bethe-Bloch formula is the inverse dependency on the particle velocity, which describes the increase in energy loss with decreasing particle velocity.

The shell correction term $\frac{C}{Z}$ becomes important only at low energy when the projectile velocity is lower than the orbital velocity of the shell electrons of the target atoms. The density correction term $\frac{\delta}{2}$ corrects for dielectric polarization effects of the traversed medium and is only important at high energies and in dense materials. Further corrections can be applied that are not included in the here presented version of the Bethe-Bloch formula. All applied corrections are very small at therapeutic energies.

1.1.3 Range of protons

The total path length R of a particle with initial energy E is given by the integration of the energy loss over the whole energy range

$$R(E) = \int_0^E \left(\frac{\mathrm{d}E'}{\mathrm{d}x}\right)^{-1} \mathrm{d}E'. \tag{1.4}$$

The quantity R is similar to the mean range of the particles since scattering is relatively small. It is assumed that the particle is continuously slowed down in many interactions while crossing the medium, hence R is referred to as *continuously slowing down approximation* (csda) range. It should be noted that the csda range is larger than the penetration depth due to scattering.
The integration of the Bethe-Bloch formula over the energy range gives the mean range of a particle beam. Due to statistical variations, particles with the same initial energy have varying path lengths. This results, for a thick absorber, in a Gaussian shaped range distribution, referred to as range straggling.

1.1.4 Sources of range uncertainties

An important part of proton therapy treatment planning is the estimation of the beam range within the patient. In an ideal scenario, it would be possible to accurately predict the beam range in human tissues. Due to uncertainties during the treatment planning and delivery process, the delivered beam range, however, differs from the predicted beam range. To ensure full coverage of the treatment volume, range uncertainties are taken into account by adding margins to the treatment volume. As a consequence, a substantial amount of healthy tissue receives the full treatment dose, increasing the risk of side effects. No clinical consensus on the size of these margins exists, but centers often use 3.5% of the predicted range plus an absolute value. The different sources of uncertainties and their contribution to the margin are summarized by Paganetti (2012) [9], and again listed here in table 1.1. Paganetti identifies two categories of sources of uncertainties in a patient treatment. The first

category of uncertainties is largely independent of the dose calculation and includes measurement uncertainties during commissioning, uncertainties from the compensator, uncertainties from beam reproducibility, and uncertainties from patient setup. The second category of uncertainties comes from the dose calculation itself and includes uncertainties from biological effects, computed tomography (CT) imaging and the conversion into relevant quantities, uncertainties from the mean excitation energy, and the uncertainties from range degradation. The main uncertainties occur in the second category, and efforts to improve range predictions should be focused on the aforementioned sources of uncertainties.

Range uncertainties from CT imaging

From a physical point of view, the reduction of uncertainties coming from dose calculation is of current research interest. Several research groups focus on the reduction of uncertainties from CT imaging and conversion to relevant quantities. In clinical practice, the range is predicted based on a single-energy CT (SECT) scan of the patient. The acquired CT numbers are converted into the quantities needed to calculate the beam range. These quantities can be relative stopping powers (RSPs, also referred to as stopping power ratio (SPR)) for analytical dose calculation techniques [10], or densities and elemental compositions of body tissues for Monte Carlo (MC) based dose calculation techniques [11]. In the first case, the conversion of CT numbers to RSP values is based on a bi-linear relation between measured CT numbers and RSP values of different tissue-equivalent plastics. The effect of nuclear interactions and secondary particles, which presents as a low dose envelope often referred to as 'beam halo', is usually estimated from a semi-empirical model [12]. In the second case, the required densities and elemental compositions are found by segmentation, i.e. assigning a range of CT numbers to a certain body tissue and assume uniform densities and compositions within the assigned tissue. From the assigned densities and compositions, RSP values are calculated for the transport of the primary beam, and nuclear interaction cross sections are calculated to simu-

late secondary particles from nuclear interactions. In both cases, errors are introduced during the conversion process since SECT does not provide enough information to accurately estimate tissue-specific quantities relevant to proton therapy planning. A popular approach for a better characterization of tissue quantities from CT imaging is to replace the currently used SECT imaging by dual-energy computed tomography (DECT) imaging. With DECT, two CT images of the same volume at different energies are acquired, providing another layer of information that can be used to estimate quantities needed for treatment planning. Different mathematical models were proposed in literature [13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23] to convert DECT images into images of RSP values that can be used as in input into parametric treatment planning algorithms. Other models focus on the estimation of densities and elemental compositions of body tissues [24, 25, 26] as input for MC-based dose calculations.

Range uncertainties from mean excitation energies

Another major source of uncertainties in dose calculation comes from the mean excitation energies, or I-values. The I-value is, next to the electron density (ED), the second tissue-specific quantity necessary to estimate the RSP from CT imaging. While the I-values of unbound elements and simple composite materials can be derived from stopping power data or calculated from semi-empirical dipole oscillator-strength distributions (table 4.1 in ICRU report 37 [27]), the I-values of human tissues carry great uncertainty. Various studies were performed in the past to accurately determine the I-value of liquid water [28, 29, 30, 31, 32, 33], with results varying by approximately 15%. These variations arise from different measurement setups, different applied corrections to the Bethe formula, as well as errors inherent to the analysis of the measurements. Considering that liquid water is a relatively simple composite material compared to human tissues, the uncertainties on human tissue I-values can be assumed similar or higher. Currently, a simple model is applied to estimate tissue I-values needed to calculate tissue RSPs, the Bragg additivity rule. With

this model, one can estimate tissue I-values by performing a weighted sum of the logarithmic elemental I-values. Associated limitations lie in the simplicity of the model since the I-value of an element changes with the bonding state of the element within a composition. Clinically used elemental I-values were proposed in 1981 [34] and adopted in the ICRU report 37 in 1984 [27], without a thorough uncertainty budget on those values or the Bragg additivity rule itself. The later ICRU report 49 [35] adapts these values for the use in proton and ion radiotherapy. Paganetti (2012) [9] recommends to account for range uncertainties from tissue I-values of 1.5%. This value is based on a paper by Andreo (2009) [36], who estimated the uncertainty on tissue I-values to be about 5-15%.

Range uncertainties from range degradation

One of the largest sources of range uncertainties in particle therapy planning arises from lateral inhomogeneities in the irradiated body. Interfaces between tissues with very different densities (such as soft tissue - lung interfaces or soft tissue - bone interfaces) introduce range degradation. Analytical dose calculation algorithms use ray tracing to calculate the beam range, based on the water equivalent thickness of human tissues in the beam path, neglecting the effects of multiple Coulomb scattering and straggling [37, 38]. This can lead to severe range uncertainties in body sites that contain inhomogeneities [39]. The review by Paganetti suggests to include a global relative range uncertainty of 2.5% to the treatment margin to account for range degradation with ray tracing techniques. This review also pointed out the potential of MC dose calculation techniques to reduce these range errors to a level of 0.1%, as demonstrated in earlier studies [40].

1.1.5 Margins in proton therapy

Uncertainties including the ones listed above limit the accuracy of radiotherapy. To ensure full target coverage, the errors and uncertainties that can arise during treatment need to be taken into account during radiotherapy plan-

Table 1.1: Sources of range uncertainties in proton therapy as listed in [9]. The here quoted values for the range uncertainties are estimates (based on 1.5 standard deviations) taken from different publications. Values express confidence intervals, those with a + and - sign refer to asymmetric intervals around the estimate. The values given in mm are absolute uncertainties, values given in % are relative uncertainties. The bottom lines are the total uncertainties, quoted as a mixture of relative and absolute uncertainties. This work focuses on the investigation of those uncertainties printed in bold.

Source of range uncertainty	without MC	with MC
Independent of dose calculation		
Measurement uncertainty in water	0.3 mm	0.3 mm
Compensator design	0.2 mm	0.2 mm
Beam reproducibility	0.2 mm	0.2 mm
Patient setup	0.7 mm	0.7 mm
Dose calculation		
Biology^	+0.8%	+0.8%
CT imaging and calibration	0.5%	0.5%
CT conversion to tissue	0.5%	0.2%
CT grid size	0.3%	0.3%
Mean excitation energies in tissues	1.5%	1.5%
Range degradation: complex inhomogeneities	−0.7%	0.1%
Range degradation: local lateral inhomogeneities*	2.5%	0.1%
Total (excluding ^,*)	2.7% + 1.2 mm	2.4% + 1.2 mm
Total (excluding ^)	4.6% + 1.2 mm	2.4% + 1.2 mm

ning. The uncertainties occurring during treatment can be of systematic or random nature. For example, errors arising from target delineation are purely systematic as they affect all fractions in the same way, whereas setup errors and errors from organ motion have a systematic and a random component [41]. Random and systematic errors have different effects on the delivered dose distribution. While Random errors blur the dose distribution [42], systematic errors introduce a shift of the cumulative dose relative to the planned dose distribution [43].

The IRCU [44] defines the treatment volumes as follows:

1. Gross Tumor Volume (GTV): "The GTV is the is the gross palpable or visible/demonstrable extent and location of malignant growth."

2. Clinical Target Volume (CTV): "The CTV is a tissue volume that contains a demonstrable GTV and/or subclinical microscopic malignant disease, which has to be eliminated. This volume thus has to be treated adequately in order to achieve the aim of therapy, cure or palliation."

3. Planning Target Volume (PTV): "The Planning Target Volume is a geometrical concept, and it is defined to select appropriate beam sizes and beam arrangements, taking into consideration the net effect of all the possible geometrical variations, in order to ensure that the prescribed dose is actually absorbed in the CTV."

In the above definitions, the PTV is an expansion of the CTV and contains all treatment uncertainties. The PTV is hence the volume to be irradiated to ensure full coverage of the CTV with the prescribed dose. However, there are no unambiguous guidelines in the above mentioned ICRU report on how to expand the CTV into a PTV. For that purpose, Stroom *et al.* [45] and van Herk *et al.* [43] provided analytical descriptions of the effect of random and systematic geometrical errors and suggest recipes on how to derive the PTV margin based on coverage probabilities. Stroom *et al.* recommend a margin of 2 times the standard deviation of the systematic errors plus 0.7 times the standard deviation of random errors. Van Herk *et al.* recommend a margin of 2.5 times the standard deviation of the systematic errors plus 0.7 times the standard deviation of random errors. The here summarized considerations were made for radiotherapy with photons and are applied in clinical routine. The principal is taken over for proton and ion radiotherapy. Additionally, the uncertainties arising from dose calculation,as listed in table 1.1, need to be considered and included in the margin recipe.

1.2 Research aims and objectives

Although it was shown in earlier studies that DECT is capable of predicting the RSP values accurately and DECT scanners are routinely available, it is

not yet used for proton therapy treatment planning. The reason for this is that the conversion models (DECT images to RSP values) were only tested in a theoretical framework using ideal CT numbers of a set of standard human reference tissues. The performance of the conversion models in realistic situations with CT artifacts and tissues is still unclear. The evaluation of the performance in realistic situations is not a straightforward task for two reasons. Firstly, it cannot be undertaken using routinely used CT calibration phantoms. These phantoms do not resemble a human body in elemental compositions and are therefore not suitable to draw conclusions on the accuracy of a conversion model. Secondly, it cannot be undertaken in patient scans, since the elemental compositions and densities of body tissues are unknown. To investigate the performance of DECT predicted RSPs in realistic situations, this project focuses on the following aims.

1. Evaluate the potential of DECT to reduce range uncertainties in a controlled theoretical environment. The objectives of this project are:

    - Implement DECT conversion models suggested in the literature.
    - Develop a virtual phantom with known ground truth and simulate DECT and SECT images containing CT artifacts and noise.
    - Compare the performance of different DECT and SECT conversion models using simulated images of the virtual phantom.
    - For the best performing DECT conversion methods, evaluate the range accuracy and uncertainties for different levels of image noise and compare to SECT.

2. Evaluate the potential of DECT to reduce range uncertainties in a realistic clinical environment. To meet this aim, the WER of animal tissues is measured in a proton beam. The measured tissue WER values give the opportunity to assess the accuracy of SECT and DECT calibrations and compare them. Furthermore, they can be used to investigate the

effect of the range calculation algorithm (RT or MC) on the predicted range. The objectives of this project are:

- Develop an efficient method to measure water equivalent ranges (WER) of a proton beam in animal tissues.
- Measure the WER of a proton beam through animal tissue samples.
- Estimate the WER based on SECT and DECT images of the samples and compare the accuracy.
- Calculate the WER based on a CT scan of the samples using a) ray tracing and b) MC to model the beam transport.
- Compare the accuracy of both models and report on the potential improvement with MC.

Since animal tissues were used in the above study, knowledge on elemental compositions, densities and therefore stopping powers is unavailable. The only measurable quantity relevant here is the WER of the proton beam after traversing the samples. When comparing this WER with those predicted from SECT or DECT, the resulting uncertainties are a combination of CT imaging/conversion uncertainties and uncertainties on the I-values. The experimentally observed range uncertainty, however, is found to be much lower than the values tabulated by Paganetti (approximately 0.7% for CT imaging/conversion and 1.5% for I-values), suggesting that these numbers are overestimated for the here investigated applications. A further limitation when estimating tissue parameters for range prediction is, therefore, the I-value itself, defining the new aim of this thesis as follows.

3. Perform a rigorous investigation of the I-value uncertainties and their influence on RSP values and beam ranges. In order to improve the understanding of I-value uncertainties and how they translate into RSP and range uncertainties, the following objectives are defined:

- Develop a theoretical model to optimize elemental I-values for the use in compounds in combination with the Bragg additivity rule.
- Establish a framework to quantify uncertainties associated with the elemental mean excitation energies and the Bragg additivity rule.
- Investigate how these uncertainties translate into RSP uncertainties and proton beam range uncertainties and provide an uncertainty budget.

Each of the above-listed aims was published as a first author paper in peer-reviewed journals during the course of this PhD.

1.3 My contribution

The work presented in this thesis was performed as a collaboration between different centers, with the involvement of scientists with clinical and theoretical backgrounds. The following centers were involved with this work:

- Proton and Advanced Radiotherapy Group, Department of Medical Physics and Biomedical Engineering, University College London (UCL), London, UK
- Acoustics and Ionising Radiation Team, National Physical Laboratory (NPL), Teddington, UK
- Department of Radiation Oncology, Massachusetts General Hospital (MGH), Boston, MA, USA
- Centre hospitalier de l'Université de Montréal (CHUM), Montréal, Québec, Canada
- Département de Physique, Université de Montréal, Montréal, Québec, Canada
- Department of Medical Radiation Physics and Nuclear Medicine, Karolinska University Hospital, Stockholm, Sweden

During the course of my PhD, I spent time at each of the first four centers. The project required work in different research areas, with my contribution to each of these areas listed below.

- **Identifying the clinical relevance of DECT for proton therapy planning:** Before starting the PhD in Medical Physics at UCL, I underwent the two years training to become a clinical scientist at the German Cancer Research Center (DKFZ) in Heidelberg, Germany. This program allowed me to spend equal amounts of time in three different facilities in Heidelberg: DKFZ, the University Clinic, and the Heidelberg Ion Beam Therapy (HIT) facility. Having acquired clinical experience in different advanced radiotherapy methods, I was well able to identify the need for better tissue characterization, especially for proton and ion radiotherapy. With the first two UK based NHS proton therapy centers at the University College London Hospital (UCLH) and the Christie in Manchester being developed, the expertise in imaging for proton therapy is very limited. After conversations with clinical scientists at UCL, UCLH, and NPL, I selected this particular project.

- **Tissue characterization methods from CT imaging:** After a thorough literature research on tissue characterization with SECT and DECT, I noticed the variety of proposed CT to RSP conversion models in the literature. Furthermore, a comparison between different models was impossible since every publication used different materials to test their models. During a discussion with my supervisors, I decided to implement and test each model on simulated images so that inter-comparison and performance evaluation was possible. I developed a method to calculate range errors and range uncertainties introduced by the CT to RSP conversion model and applied it to each method.

- **Development of experimental range measurement techniques:** I worked together with scientists at NPL and MGH to develop an efficient

method to measure water equivalent beam ranges in animal tissue samples. I spent a full year at the Francis H. Burr proton therapy center in MGH to help developing an efficient and accurate measurement method. I was involved in the development and calibration of the method. I tested the developed method using water phantoms and different tissue-equivalent materials.

The range measurement technique was used to verify the proton ranges predicted by SECT and DECT in animal tissue samples. For this part of the project, I performed the experiments and CT scans and wrote the codes to convert CT numbers into RSP values. I adapted an existing ray tracing algorithm to predict the WER from the RSP values of the tissue samples.

- **MC treatment planning:** When predicting beam ranges from SECT or DECT, one first voxelwise converts CT numbers to RSP values or elemental compositions using a conversion model. During the analysis of the experimental data, I first decided to use a simple ray tracing algorithm to predict the beam ranges from the RSP maps, as it is commonly done in clinically used treatment planning systems. The observed errors, especially in samples containing bone tissues, were larger than expected. I decided to use MC to model the beam transport through my samples, which largely improved the range prediction. Since measured WER values for those samples were available, I decided that it was feasible to investigate the range prediction errors introduced by ray tracing as compared to MC.

- **Quantification of range uncertainties:** As above mentioned, I had quantified the range uncertainties arising from the CT to RSP conversion model on a theoretical ground. After the analysis of the animal tissue experiments, I noticed that the observed uncertainty is much lower than the uncertainty quoted by Paganetti (2012). Since the observed uncer-

tainty is a combination of uncertainties from CT to RSP conversion and I-value, I decided to work on the quantification of I-value uncertainties alone. I developed a mathematical model to optimize the I-values for the use in compound materials and estimate their uncertainties. I propagated the uncertainties to RSP and range uncertainties, such that a better understanding of range uncertainties stemming from I-value uncertainties alone is possible.

1.4 Impact and novelty of the work

The research presented in this thesis focuses on the improvements in range accuracy and uncertainty achievable with DECT. The following aspects are novel:

- **Comparing the performance of various DECT conversion models:** In recent literature, many models were published to convert DECT images into maps of RSP values or elemental compositions. These papers quote the accuracy of the proposed model in a self-consistency manner, meaning that the model is tested on the same data as they were calibrated on. Furthermore, each published method uses different calibration materials, making the comparison of accuracies between models difficult. This thesis provides the first comparison of different suggested models and their accuracies to predict proton ranges. This part of the work impacts clinical decision making, by giving clinicians an overview of DECT conversion models proposed in literature, their accuracy and the uncertainty related to the conversion model.

- **Sensitivity of DECT predicted tissue parameters to image noise:** Since the published models were only tested in a theoretical scope, it was unclear how sensitive the predicted RSP and range values are to imaging artifacts. This thesis presents the first investigation of the influence of noise and beam hardening on these models. An important finding of this study is that noise impacts the DECT to RSP

conversion, to a point where the benefits of using DECT can be lost over the noise robustness of SECT conversion models. This important result impacts future work, therefore efforts to make DECT conversion models more robust to noise were made hereafter, such as methods suggested in Bär *et al.* (2018) [46] or Lalonde *et al.* (2017) [47].

- **Modeling of range uncertainties:** The developed theoretical method to assess range accuracy and uncertainty from RSP errors allows the quantification of range uncertainties introduced from the CT to RSP conversion method. This method was used in Lalonde *et al.* (2017) [47] to assess the range accuracy and uncertainties when using multi-energy CT (MECT) to estimate elemental compositions. This method is also applicable to other RSP prediction or measurement methods, such as proton or ion tomography.

- **Validation of DECT predicted ranges using animal tissue samples:** This thesis presents the first end-to-end test of range predictions with DECT and compare them to SECT in animal tissue samples. It is the first to demonstrate the achievable improvement in range uncertainty and the first to demonstrate the benefits of DECT in a clinical setup. This part of the work states the superiority of DECT over SECT to predict RSPs and elemental compositions in body tissues and therefore impacts the future of radiotherapy planning strategies. This study shows that DECT can safely be used for proton and ion therapy planning, and might lead to an adaption of clinical margins.

- **Experimental investigation of range errors introduced by the dose calculation algorithm:** This thesis presents the first study to demonstrate the range uncertainties introduced by a ray tracing algorithm in a situation where the ground truth is known, and the improvements that can be achieved by using MC for range prediction. It provides further rationale to push for full MC treatment planning systems. Ad-

ditionally, it provides an estimate of the range uncertainties introduces when ray tracing is used.

- **A new set of elemental *I*-values and uncertainties:** This thesis revisits elemental I-values since the clinically applied values were established in 1981 and not revisited since. The here proposed elemental I-values were established using a mathematical model and a large set of calibration data. This thesis also presents the first thorough and rigorous model to estimate range uncertainties coming from I-values alone. The demonstrated uncertainties can be used clinically to adjust the treatment margin.

Chapter 2

The potential of dual-energy CT to reduce proton beam range uncertainties

The work presented in this chapter can be found in the following peer-reviewed journal articles:

1. **Bär, E.**, Lalonde, A., Royle, G., Lu, H.-M. and Bouchard, H. (2017). The potential of dual-energy CT to reduce proton beam range uncertainties. *Medical Physics* 44: 2332-2344. doi:10.1002/mp.12215.

2. Lalonde, A., **Bär, E.** and Bouchard, H. (2017) A Bayesian approach to solve proton stopping powers from noisy multienergy CT data. *Medical Physics* 44: 5293-5302. doi: 10.1002/mp.12489.

It was also presented at the following conferences:

3. **Bär, E.**, Lalonde, A., Royle, G., Lu, H.-M. and Bouchard, H. (2016). On the Performance of Four Dual Energy CT Formalisms for Extracting Proton Stopping Powers. Washington D.C., *58th Annual Meeting of the AAPM. Medical physics* 43 (6Part11): 3453-3453. doi: 10.1118/1.4956103.

4. **Bär, E.**, Lalonde, A., Royle, G., and Bouchard, H. (2016). On the

performance of dual-energy CT for determining proton stopping powers. London, *3rd PPRIG workshop*.

Contribution of Authors: Arthur Lalonde developed the codes for the SECT and DECT elemental decomposition methods, *i.e.* the method by Schneider *et al.* (2000) for SECT, and the methods by Landry *et al.* #2, Hünemohr *et al.* #2 and Lalonde and Bouchard for DECT. Gary Royle and Hsiao-Ming Lu provided valuable input concerning the clinical importance and impact of this work. Hugo Bouchard supervised the project and guided the development of ideas, methods, results, and conclusions of this work. All other work presented in this chapter was done by myself.

2.1 Introduction

Conventionally, radiotherapy planning is based on CT images. For proton therapy dose calculation, CT numbers need to be converted into tissue SPRs relative to water, which are used to calculate the beam range in the patient and the energy deposited along the penetration path. To exploit the full benefits of protons and to avoid errors on dose delivery at the distal fall-off, accurate conversion from CT numbers to SPR is essential. To further improve clinical outcomes of proton therapy, one must aim at a higher precision, which allows us to reduce safety margins and thus irradiate less healthy tissue while maintaining conformal target dose.

In clinical practice, human tissue characterization for treatment planning is achieved by acquiring a CT scan on the patient and then converting the data into SPRs. Conventionally, the CT scan is acquired using a single energy spectrum, and one clinically reliable method to obtain SPR from CT numbers is the calibration method proposed by Schneider *et al.* (1996) [10]. In this procedure, a relation between calculated SPRs of human reference tissues [48, 49], and CT numbers (in Hounsfield units (HU)) are determined using a plastic phantom with radiological properties equivalent to that of human tissues.

The calibration of Schneider *et al.* (1996), referred to as the SECT stoichiometric calibration method throughout this chapter, is fairly accurate in predicting human tissue SPR [50]. Schaffner and Pedroni (1998) verified the SECT stoichiometric calibration by measuring pairs of CT numbers and SPR using animal tissue samples. They found an accuracy in SPRs of ± 1.1% for soft tissues and ± 1.8% for bones, which translates into range uncertainties of up to 3 mm for therapeutic energies. In more recent work, the combined uncertainty in proton range estimation was reported to be 2.7-3.5% + 1.0-1.2 mm (1.5 standard deviations), excluding biological effects [51, 9]. While a large uncertainty is associated with the knowledge of the mean excitation energy (I-value) [9], another limitation in the accuracy of proton beam treatment planning is from CT data. In SECT, data is limited to a single dimension per voxel and this is problematic since HU-SPR calibration curves are not one-to-one relations (i.e., bijections) for human tissues. While both HU and SPR values are dominated by the ED, these quantities depend on other tissue properties, such as the effective atomic number (EAN) or the I-value [51, 9]. In turn, these properties depend on the elemental composition. Small patient-to-patient variations in density and elemental compositions were shown to introduce significant changes in CT numbers [23]. These variations are not necessarily resolved by the SECT stoichiometric calibration since the HU-to-SPR conversion approach cannot explicitly decouple the dependency of CT numbers on elemental compositions and mass density, therefore limiting the precision to which tissue characteristics can be resolved.

DECT has the potential to improve the conversion of CT data to SPR. Over the last decade, several papers were published on DECT to either show potential benefits for radiotherapy or to propose a mathematical formalism to extract tissue parameters relevant to dose calculation. Recent publications propose the extraction of ED and EAN (or alternatively, extraction of the I-value), from DECT images [52, 53, 21, 16, 15, 13, 20, 14, 26]. These methods rely on post-reconstruction data analysis, conversely to sinogram-based

methods (e.g., Refs. [54, 55]) which are yet to be fully explored. Studies on DECT for proton therapy typically report errors on stopping power determination between 0.5% and 1.5% [52, 53, 21, 16, 15, 13, 20, 14, 26]. Although there exists no direct relation between X-ray attenuation and stopping powers, it was shown that DECT has the potential to substantially improve proton radiotherapy planning as it is widely clinically available.

This study is aimed at evaluating the potential of DECT to reduce proton beam range uncertainties in a clinical context, with focus on CT artifacts and noise, and leaving aside uncertainties related to the I-value and the CT grid size. The performance of different mathematical techniques to predict proton stopping powers is compared theoretically, experimentally and with simulated CT data. Since a consistent basis is needed for such comparison, all methods are compared under the same conditions using the same calibration phantom, CT images, and statistical quantities. The resulting distributions of SPR errors are used to estimate the impact on proton beam range uncertainties, in this way allowing estimating the gain in precision provided by DECT in a clinical environment. A comparison against the SECT stoichiometric method is achieved in order to predict the potential clinical impact of DECT in proton therapy dose calculation.

2.2 Methods

2.2.1 Photon attenuation in mixtures and compounds

The attenuation of photons traveling through media is described by the Lambert-Beer law

$$\phi(x) = \phi(0) \cdot \mathrm{e}^{-\mu x}, \tag{2.1}$$

where $\phi(x)$ is the photon flux at point x, $\phi(0)$ is the initial flux and is μ the linear attenuation coefficient. The linear attenuation coefficient is defined as

$$\mu = \begin{cases} \rho \frac{Z}{A} N_A \sigma_e & \text{for electronic interactions} \\ \rho \frac{1}{A} N_A \sigma_a & \text{for atomic interactions,} \end{cases} \tag{2.2}$$

where ρ is the mass density of the material Z is the atomic number, A is the mass number, N_A is Avogadro's number and $\sigma_{e,a}$ is the interaction cross section for electronic or atomic interactions. Generally, the interaction cross section σ is dependent on the atomic number Z of the interacting medium and the photon energy E,

$$\sigma \equiv \sigma(Z, E). \tag{2.3}$$

The electron density of a medium is given by

$$N_{\mathrm{e}} = \rho N_{\mathrm{A}} \left(\frac{Z}{A}\right)_{\mathrm{med}}. \tag{2.4}$$

We define the average $\left(\frac{Z}{A}\right)_{\mathrm{med}}$ as

$$\left(\frac{Z}{A}\right)_{\mathrm{med}} = \sum_i \omega_i \left(\frac{Z}{A}\right)_i, \tag{2.5}$$

with ω_i being the fraction of mass associated with the i-th element.

The mass attenuation coefficient for a mixture or compound can be approximated as the weighted sum of attenuation coefficients of the constituents i [56]

$$\begin{aligned} \left(\frac{\mu}{\rho}\right)_{\mathrm{med}} &= \sum_i \omega_i \frac{\mu_i}{\rho_i} \\ &= N_A \left(\frac{Z}{A}\right)_{\mathrm{med}} \sigma_{\mathrm{e,med}}. \end{aligned} \tag{2.6}$$

This expression neglects any changes in attenuation properties resulting from molecular binding.

In CT imaging, the attenuation coefficient is measured in terms of CT numbers, expressed in Hounsfield Units (HU):

$$\mathrm{CT\,number} = \frac{\mu - \mu_{\mathrm{w}}}{\mu_{\mathrm{w}}} \cdot 1000, \tag{2.7}$$

with μ being the attenuation coefficient of the respective medium at the re-

spective energy and μ_{w} being the attenuation coefficient of water. From CT numbers, the tissue attenuation coefficient can be derived.

From eq. (2.2) we know that the attenuation coefficient μ of a medium is defined as the product of the electron density relative to water ρ_e and the electronic cross section σ_{e}. The electronic cross section describes the probability of a photon interacting with the material. It can therefore be expressed as the sum of the probabilities of the interactions

$$\sigma_{\mathrm{e}}(E,Z) = \frac{\sigma_{\mathrm{a}}(E,Z)}{Z} = \left[\sigma_{\mathrm{ph}}(E,Z) + \sigma_{\mathrm{incoh}}(E,Z) + \sigma_{\mathrm{coh}}(E,Z)\right]_{\mathrm{e}}, \qquad (2.8)$$

where σ_{ph} is the cross section of the photoelectric interaction, σ_{coh} is the cross section of coherent scatter and σ_{incoh} is the cross section of incoherent scatter. The interaction cross section is a function of the photon energy E and the atomic number Z of the element involved in the interaction. If the interaction happens with a medium or compound, the atomic number is replaced by the EAN of the medium. In literature, several parametrization of $\sigma(E,Z)$ have been proposed.

- **Alvarez and Macovski, 1976**

  Alvarez and Macovski [57] suggested to parametrize the attenuation coefficient in the form of

  $$\mu(E,Z) = N_{\mathrm{e}}\left(K_1\frac{Z^k}{E^l} + K_2 f_{\mathrm{KN}}(E)\right), \qquad (2.9)$$

  with constants K_1 and K_2 and the Klein-Nishina function $f_{\mathrm{KN}}(E)$. The exponents k and l are energy dependent, but classically approximated with $k = 3$ and $l = 3$. The first term describes the energy dependence of photoelectric interactions, which is approximated as $\frac{Z^3}{E^3}$. The second term gives the energy dependence of the total Compton scattering cross sections, as derived by Klein and Nishina in 1929 [58]. The influence of coherent scatter is not considered in this parametrization.

- **Rutherford *et al.*, 1976**

  Rutherford [59] suggested a general expression for the parametrization of cross sections, which is given as

$$\sigma = AE^{-m}Z^n. \tag{2.10}$$

  The parameters A, m and n were found via fitting the parameters to cross section data of oxygen. The resulting expression for the attenuation coefficient is as follows

$$\mu = 20.64E^{-3.82}Z_{\text{eff}}^{4.62}N_{\text{eff}} + f_{\text{KN}}Z_{\text{eff}}N_{\text{eff}} + 2.80E^{-2.02}Z_{\text{eff}}^{2.86}N_{\text{eff}}, \tag{2.11}$$

  with N_{eff} being the effective number of atoms per volume and $N_{\text{e}} = Z_{\text{eff}}N_{\text{eff}}$. The first term describes attenuation via photoelectric absorption, the second term is the Klein-Nishina cross section and the third term corrects for coherent scatter and binding energy effects.
  In a more general way and not confined to the cross section data of oxygen, eq. (2.11) can be written as follows

$$\mu = N_{\text{e}}\left[K^{\text{ph}}Z^m + K^{\text{coh}}Z^n + K^{\text{KN}}\right], \tag{2.12}$$

  with the Klein-Nishina cross section [58] K^{KN}, m and n are energy dependent.

- **Jackson and Hawkes, 1981**

  Jackson and Hawkes [60] describe the cross section based on the formal theory of photoelectric effect and scattering. Thereby, the atomic cross section of the photoelectric effect is given as

$$\begin{aligned} {}_{\text{a}}\sigma^{\text{pe}}(E,Z) &= \left[4\sqrt{2}Z^5\alpha^4\left(\frac{mc^2}{E}\right)^{\frac{7}{2}}\frac{8}{3}\pi r_e^2\right] \\ &\times\left[2\pi\left(\frac{\epsilon_{\text{K}}}{E}\right)^{\frac{1}{2}}f(n1)\right]\left[1+F_{\text{ns}}(\beta)\right]\text{UN}(E,Z). \end{aligned} \tag{2.13}$$

In the first factor, α is the fine structure constant, mc^2 is the electron rest mass energy, r_e^2 is the classical electron radius. In the second factor, ϵ_{K} describes the k-shell binding energy and is given as

$$\epsilon_{\mathrm{K}} = \frac{1}{2}(Z\alpha)^2(mc^2), \tag{2.14}$$

and

$$n_1 = \left[\frac{\epsilon_{\mathrm{K}}}{(E-\epsilon_{\mathrm{K}})}\right]^{\frac{1}{2}}, \tag{2.15}$$

$$f(n_1) = \frac{\exp(-4n_1\cot^{-1}(n_1))}{1-\exp(-2\pi n_1)}. \tag{2.16}$$

The first two factors in eq. (2.13) represent the atomic Stobbe cross section (${}_{\mathrm{a}}\sigma_{\mathrm{ls}}^{\mathrm{ST}}$) for the k-shell. The third factor is a correction factor accounting for relativistic effects, given as

$$[1+F_{\mathrm{ns}}(\beta)] = 1 + 0.143\beta^2 + 1.667\beta^8, \tag{2.17}$$

with $\beta = \frac{v}{c}$, with v is the velocity of the photoelectron. The last factor $\mathrm{UN}(E,Z)$ accounts for screening of the nucleus by atomic electrons.
The combined cross section for coherent and incoherent scattering is proposed as

$$\begin{aligned} {}_{\mathrm{a}}\sigma^{\mathrm{coh}}(E,Z) +_{\mathrm{a}} \sigma^{\mathrm{incoh}}(E,Z) = \\ Z\,{}_{\mathrm{e}}\sigma^{\mathrm{KN}}(E) + (1-f(Z)Z^{-1})\left[\left(\frac{Z}{Z_{\mathrm{s}}}\right)^2_{\mathrm{a}}\sigma^{\mathrm{coh}}(E_{\mathrm{s}},Z_{\mathrm{s}})\right], \end{aligned} \tag{2.18}$$

with ${}_{\mathrm{e}}\sigma^{\mathrm{KN}}(E)$ the Klein-Nishina cross section, $f(Z) = Z^{0.5}$, Z_{s} is a standard element used for scaling the coherent scatter cross section as a function of the standard energy $E_{\mathrm{s}} = (Z_{\mathrm{s}}/Z)^{1/3}E$. Jackson and Hawkes suggested to use oxygen as a standard element for human tissues. The

Klein-Nishina cross section is calculated as

$$
\begin{aligned}
{}_{\mathrm{e}}\sigma^{\mathrm{KN}}(E) &= 2\pi r_e^2 \\
\times \left\{ \frac{1+\delta}{\delta^2} \left[\frac{2(1+\delta)}{1+2\delta} - \frac{1}{\delta} \ln(1+2\delta) \right] + \frac{1}{2\delta} \ln(1+2\delta) - \frac{1+3\delta}{(1+2\delta)^2} \right\},
\end{aligned}
\tag{2.19}
$$

with $\delta = E/mc^2$. The total cross section σ^{tot} is calculated using eq. (2.8).

- **Torikoshi *et al.*, 2003**

  Torikoshi *et al.* [61] simplified eq. (2.8) to

$$
\mu = N_{\mathrm{e}} \left[Z^4 F(E,Z) + G(E,Z) \right], \tag{2.20}
$$

  where the first term approximates photoelectric interactions and the second term summarizes the scatter contribution. In their paper, they measured the electron density of tissues with monochromatic synchrotron radiation. For $F(E,Z)$ and $G(E,Z)$ they used the physics-based formulas proposed by Jackson and Hawkes [60]. Bazalova *et al.* [53] extended this approach for the application in spectral CT. In this approach, the functions $F(E,Z)$ and $G(E,Z)$ are obtained by quadratic fits of the photoelectric and scattering terms of XCOM [62] attenuation coefficients.

- **Bourque *et al.*, 2014**

  Bourque *et al.* [13] proposed to parametrize the electronic cross section as a function of the atomic number

$$
\sigma_{\mathrm{e}} = \sum_{m=1}^{M} b_m Z^{m-1}. \tag{2.21}
$$

  The parameters a_m are obtained by a least square fit on cross section data from the XCOM database, averaged over a given spectrum. M is the order of the fit.

This last section provided a detailed view on the cross section parametrization proposed in literature. The parametrization provided by Alvarez and Macovski and Rutherford *et al.* rely on a power law with an energy dependent exponent. In these cases, optimization is necessary to find the exponents suitable for the energy spectrum used in specific studies. The use of the classical values is not sufficiently accurate in most of the cases and will introduce uncertainties to the attenuation coefficient. The parametrization of Jackson and Hawkes and Torikoshi *et al.* in turn incorporate the physical equations related to the interactions of radiation with matter. These equations are energy dependent and therefore more suitable to describe a range of different experimental setups. However the implementation of these formulas is complex and might not be practical. The parametrization provided by Bourque *et al.* consists of a practical Taylor expansion, where parameters are found by fitting to experimentally determined cross section values. Due to the high number of degrees of freedom, it can accurately depict cross sections and thus attenuation coefficients.

2.2.2 An overview of investigated DECT tissue characterization methods

In literature, several techniques to extract proton stopping powers from DECT images exist. Commonly, these methods extract the density ρ, or alternatively, the relative electron density ρ_{e}, plus the EAN Z_{eff} [63] or Z_{med} [13] to derive the I-value via a parametric relationship between Z and I for human tissues [23, 13]. Some published methods [14, 26] do not require the concept of EAN to determine tissue parameters. The key elements of all the formalisms studied are summarized in tables 2.1 and 2.2.

While they are reported in chronological order, there are two types of techniques compared. The first type is based on parameter extraction, i.e., either ρ_{e}-Z or ρ_{e}-I. With this type, proton SPRs can be calculated with Bethe's equation:

$$S = \rho_{\mathrm{e}} \frac{k_0}{\beta^2} \left[\ln \left(\frac{2 m_e c^2 \beta^2}{I(1-\beta^2)} \right) - \beta^2 \right], \tag{2.22}$$

by taking the ratio of the resulting stopping power S for a given ρ_e, I-value and reference energy. Note that in this chapter, $I_w = 74.2$ eV is used to calculate the theoretical stopping powers, as derived from the effective atomic number of water, see table 1 in Bourque *et al.* [13]. It should be noted that changing I_w in this study will not influence the results as long as it is kept consistent. In the equation above, k_0 is a constant, m_{e} is the electron rest mass, c is the speed of light, and β the particle velocity relative to c. For the techniques extracting Z, I is calculated depending on which definition of the EAN is applied. The conversion Z into I proposed either by Yang *et al.* [23] or Bourque *et al.* [13] is used for Z_{eff} or Z_{med}, respectively.

The second type of technique is meant to extract elemental weights fractions and mass or electron density. There exists three methods compared herein predicting elemental weight fractions and density from DECT. From the predictions of these methods, the I-value of each pixel is calculated using the Bragg additivity rule [27]:

$$\ln I = \sum_i \lambda_i \ln I_i. \tag{2.23}$$

where λ_i are the elemental electronic fractions. The SPR is then obtained with equation 2.22 using the electron density, either obtained directly or calculated from the mass density and the elemental composition allowing estimating Z/A. It is worth noting that methods predicting elemental compositions and density are suitable with MC radiation transport algorithms, which are known to improve the accuracy of range predictions in heterogeneous media [9]. However, since most clinical dose calculation engines require SPRs, the present focus is on the ability to predict these ratios and further evaluate the impact on beam range predictions using an analytic model.

2.2.2.1 Bazalova *et al.* 2008

A tissue characterization method for monoenergetic photons was proposed by Torikoshi *et al.* [61], but first adapted by Bazalova *et al.* [53] for the use in commercial CT scanners. In this parametrization, the photoelectric attenuation and Compton scattering are expressed as quadratic functions $F(E, Z)$ and $G(E, Z)$. $F(E, Z)$ and $G(E, Z)$ are obtained by fitting of quadratic functions to elemental cross sections (i.e., the XCOM database [62]). For the use of this parametrization in a spectrum of energies, spectral weights and integration over the energy must be taken into account. Z_{eff} is found via numerical solution from two energies, ρ_{e} is obtained by substitution of Z_{eff}. In Bazalova *et al.*'s method presented here, the numerical solution for Z_{eff} is obtained using the MATLAB (The MathWorks, Inc., Natick, MA, USA) build in numerical solver *fzero*. Additionally, spectral attenuation in the examined object must be taken into account. Hence, the output spectrum of the X-ray tube is not used for tissue parameter extraction, but a tissue filtered spectrum. This tissue filtered spectrum is calculated using an analytical absorption model, which employs the attenuation law. As Bazalova *et al.* evaluated in their paper, it is valid to assume a filtering of 16 cm of water to describe every position within the round-shaped phantom.

2.2.2.2 Landry *et al.* 2013

To extract the EAN with DECT, Landry *et al.* [16] developed a method combining previously proposed techniques. The approach was inspired by the SECT stoichiometric calibration by Schneider *et al.* (1996) [10]. The parametrization of Rutherford *et al.* [59] was utilized. This parametrization, in contrast to the parametrization by Alvarez and Macovski, comprises a term to take coherent scatter into account. In their method, Landry *et al.* proposed using the ratio of attenuation coefficients measured with the CT scanner at low and high energy in a two-step calibration procedure. In a first step, the attenuation coefficients of a calibration phantom are measured at two energy

spectra. The measured values are used to find the stoichiometric parameters $k_{1\mathrm{kVp}}$ and $k_{2\mathrm{kVp}}$ as proposed by Schneider *et al.* (2000) [11] per energy. These parameters are then used to calculate attenuation coefficients of a set of human reference tissues [49]. The ratio of the calculated attenuation coefficients of human tissues serves as a basis data set to find the fit parameters $A_{\mathrm{l,h}}$, $B_{\mathrm{l,h}}$ and $C_{\mathrm{l,h}}$. These parameters correspond to A, B and C in table 2.1, with l for the low and h for the high energy spectrum. Z_{eff} is obtained by solving the parametrization for Z. To determine the electron density, Landry *et al.* recommended that the method by Saito is used to obtain ρ_{e}. Saito [21] developed a method to only extract electron densities from DECT. This approach employs a ΔHU, which is obtained as a linear combination of $\mathrm{HU_l}$ and $\mathrm{HU_h}$, with a single weighting factor. This factor is scanner-specific and must be found in a calibration process, employing a calibration phantom.

As an extension of their method, Landry *et al.* [25] proposed a segmentation method to extract a full elemental composition from any Z_{eff} and ρ_{e} couple. First, these two parameters are calculated for a dataset of reference human tissues. Then, the tissue assigned in each voxel is the one showing the shortest generalized distance with the measured data in the ρ_{e}-Z_{eff} space. The segmentation technique allows assigning a tissue to each voxel and a generic elemental composition to the tissue. The determination of electron density and EAN with the method of Landry *et al.* is referred to as Landry *et al.* # 1, while the one extracting elemental weights is Landry *et al.* # 2.

2.2.2.3 Hünemohr *et al.* 2014

The first of the existing DECT tissue parameter extraction methods for clinical use was published in 2003 by Heismann *et al.* [52]. They employed the attenuation cross section (μ) parametrization from Alvarez and Macovski [57] and developed their formalism on post-reconstruction data. In the model, one first term describes the attenuation due to the photoelectric effect, while the other term describes Compton scattering. Each physical effect has an associated coefficient (α and β) which quantifies the magnitude of the effect. The coeffi-

cients are energy-specific and can be found in a calibration process employing a DECT scan of materials with known compositions. The energy dependence of the system is furthermore taken into account by introducing parameters (g_{L} and g_{H}) and integrating the attenuation coefficient over the energy spectrum using spectral weights $w_{\mathrm{L,H}}$. Hünemohr *et al.* [15] adapted the approach by Heismann *et al.* and proposed a calibration using a tissue characterization phantom instead of the integration over the spectral energies. Furthermore, the authors employ the mathematical methodologies of the ρ-Z projection of Heismann *et al.* but substitute the mass density ρ by the relative electron density ρ_{e}. In this work, we chose to implement the version of Hünemohr *et al.*, employing ρ_{e} instead of ρ.

To take into account potential elemental composition variation for a given tissue within a population, Hünemohr *et al.* [24] proposed to parametrize elemental weights as a function of Z_{eff} and ρ_{e}. Thus, for each of the 13 elements (H, C, N, O, Na, Mg, P, S, Cl, K, Ca, Fe and I), a reference dataset of tissues is used to create a linear fit describing the weight of each element as a combination of ρ_e, Z_{eff}, and $\rho_e Z_{\mathrm{eff}}$, as recommended in their publication. In the present work, the determination of electron density and EAN with the method of Hünemohr *et al.* is referred to as Hünemohr *et al.* # 1, while the one allowing to obtain elemental weights is Hünemohr *et al.* # 2.

2.2.2.4 Bourque *et al.* 2014

In the method by Bourque *et al.* [13], the attenuation coefficient relative to water is parametrized as a polynomial of the order M-1 with coefficients b_m. The parameters b_m are obtained from a least square fit to measured μ/μ_{w} from a CT scan of the calibration phantom. A specific definition of the EAN is used, Z_{med}, and their values for the phantom materials have previously been calculated and averaged for both energy spectra. The fit procedure to obtain coefficients b_m must be performed for both energies of the DECT scan separately. In analogy to the attenuation coefficient, Bourque *et al.* define a parametrization for the estimation of the EAN, as listed in table 2.1. It

uses the dual-energy ratio Γ (defined as the attenuation coefficient of the low-energy scan relative to the high-energy scan) for its independence on electron density. To find the model parameters c_k, Γ is measured for the inserts of the calibration phantom and a least square fit of order $K-1$ is performed. For a dual energy CT scan of unknown tissues, Z_{med} and ρ_{e} are found by measurement of $(\mu/\mu_{\mathrm{w}})_{\mathrm{L}}$ and $(\mu/\mu_{\mathrm{w}})_{\mathrm{H}}$.

2.2.2.5 Van Abbema *et al.* 2015

Van Abbema *et al.* [20] developed a method that is not based on calibration but requires spectral knowledge. They use the electron cross-section parametrization ${}_{\mathrm{e}}\sigma^{\mathrm{tot}}(E,\widehat{Z})$ of Jackson and Hawkes [60], extended with fit functions to yield a dependency on E and Z. Knowledge of the spectral weighting function $w(E)$ at every energy increment $\mathrm{d}E$ is necessary. Z_{eff} is found by solving the ratio of attenuation coefficients at low and high energy numerically for Z, and ρ_{e} is obtained by substitution of Z_{eff}. As this method makes use of spectral knowledge, the attenuation of the examined object must be taken into account, similarly to the method proposed by Bazalova *et al.*. To account for spectral hardening, van Abbema *et al.* propose to apply a $w(E)$ local weighting function (LWF), which is obtained iteratively from spectral weights $w(E)$ and the measured attenuation coefficients in the corresponding voxel.

2.2.2.6 Han *et al.* 2016

A recent paper by Han *et al.* [14] proposed a two-parameter model. They assume that the attenuation coefficient of an unknown material in a given voxel can be described as a linear combination of the attenuation coefficient of two basis materials μ_1 and μ_2. The basis materials are chosen as water and polystyrene for soft tissues, and water and an aqueous $CaCl_2$ solution (23%) for bony tissues. The parameters c_1 and c_2 are material specific, found by measuring the attenuation coefficients of the basis materials as well as the unknown material at two different energies. The integration over all energies of the spectrum is approximated in this model by using the mean energy of

spectrum. ρ_e and I of unknown tissues are then found using the determined parameters c_1 and c_2, according to table 2.1.

2.2.2.7 Lalonde and Bouchard 2016

Lalonde and Bouchard [26] introduced a representation of human tissues based on principal component analysis (PCA). An optimal basis of virtual materials (principal components, PC) is defined from a reference dataset of tissues, each of them described by a mass density and array of elemental compositions (H, C, N, O, Na, Mg, P, S, Cl, K, Ca, Fe and I). The partial electronic density y_k of each PC is retrieved by performing a material decomposition from DECT data. Once the y_k are solved, their sum equals the electronic density and the elemental composition is unfold from the PC content. To estimate the electronic cross section of each PC (i.e., f_k in table 2.1), a calibration method similar to Bourque *et al.* [13] is proposed, but without the need for defining the EAN. In this way, the attenuation coefficient relative to water is parametrized using a series of power specific average atomic numbers, i.e., $\overline{Z}, \overline{Z^2}, \overline{Z^3}, \ldots$, referred to as Z-space. The fit parameters are obtained for each energy and scanning protocol from a least square fit on measured μ/μ_w from a CT scan of a calibration phantom. It should be noted that only the formalism of Lalonde and Bouchard gives directly a complete set of elemental weights and mass density without an intermediate step. However, two other methods (Landry *et al.* # 2 and Hünemohr *et al.* # 2) can be adapted to convert measured ρ_e and Z to suitable MC inputs. These methods are investigated in this study and compared to the PCA approach of Lalonde and Bouchard.

2.2.3 Comparison of DECT tissue characterization methods

This section describes how the performance of the different DECT methods is compared. Firstly, a theoretical comparison with the XCOM photon cross sections database is performed in order to evaluate the theoretical robustness

Table 2.1: Summary of the theoretical foundation of different DECT formalisms.

	μ parametrization	Z definition	Requires CT calibration
Bazalova *et al.*	$\mu = N_\mathrm{e} \sum_i w_i \left(Z^4 F(E_i, Z) + G(E_i, Z)\right)$	Mayneord ($m = 3.5$)	No
Landry *et al.* #1 and #2	$\mu = N_\mathrm{e} \left(A + BZ^m + CZ^n\right)$	Mayneord ($m = 3.3$)	Yes
Hünemohr *et al.* #1 and #2	$\mu = N_\mathrm{e} \left(\alpha \frac{Z^m}{E^t} + \beta\right)$	Mayneord ($m = 3.1$)	Yes
Bourque *et al.*	$\mu/\mu_\mathrm{w} = \rho_\mathrm{e} \sum_{m=1}^{M} b_m Z^{m-1}$	Behavior of electronic cross sections for elements	Yes
Van Abbema *et al.*	$\mu = N_\mathrm{e} \int_0^\infty w(E)\, {}_\mathrm{e}\sigma^\mathrm{tot}(E, \widehat{Z})\, \mathrm{d}E$	Behavior of $\frac{\mu_\mathrm{L}}{\mu_\mathrm{H}}$ for mixtures	No
Han *et al.*	$\mu = c_1\mu_1 + c_2\mu_2$	None	Yes
Lalonde and Bouchard	$\mu/\mu_\mathrm{w} = \overline{y}_0 f_0 + \sum_{k=1}^{K} y_k f_k$	None	Yes

Table 2.2: Summary of different formalisms to predict tissue parameters with DECT.

	EAN	I-value	ED
Bazalova *et al.*	solve $\frac{u_\mathrm{L}}{u_\mathrm{H}}$ numerically	Yang *et al.*	substitute $\widehat{Z}$
Landry *et al.* #1 and #2	solve $\frac{u_\mathrm{L}}{u_\mathrm{H}}$ for Z	Yang *et al.* Bragg additivity rule	$\widehat{\rho}_\mathrm{e} = \frac{\Delta\mathrm{HU}}{1000} + 1$
Hünemohr *et al.* #1 and #2	substitute $\widehat{\rho}_\mathrm{e}$	Yang *et al.* Bragg additivity rule	$\widehat{\rho}_\mathrm{e} = \frac{1}{\beta}\frac{g_\mathrm{L}\mu_\mathrm{H} - g_\mathrm{H}\mu_\mathrm{L}}{g_\mathrm{L} - g_\mathrm{H}}$
Bourque *et al.*	$\widehat{Z}_\mathrm{eff} = \sum_{k=1}^{K} c_k \Gamma^{k-1}$	5^th-order fit with Z_med	$\widehat{\rho}_\mathrm{e,L/H} = \frac{u_\mathrm{L/H}}{\sum_{m=1}^{M} b_{m,\mathrm{L/H}} Z_\mathrm{eff}^{m-1}}$
Van Abbema *et al.*	solve $\frac{\mu_\mathrm{L}}{\mu_\mathrm{H}}$ numerically	Yang *et al.*	substitute $\widehat{Z}$
Han *et al.*	None	$\widehat{I}_x = f_I \left(\frac{c_1}{c_1+c_2}\right) \exp\left(\frac{c_1\rho_\mathrm{e1}\ln(I_1) + c_2\rho_\mathrm{e2}\ln(I_2)}{c_1\rho_\mathrm{e1} + c_2\rho_\mathrm{e2}}\right)$	$\widehat{\rho}_{ex} = c_1\rho_\mathrm{e1} + c_2\rho_\mathrm{e2}$
Lalonde and Bouchard	None	Bragg additivity rule	$\widehat{\rho}_\mathrm{e} = \overline{y}_0 + \sum_{k=0}^{K} y_k$

of the method. Secondly, methods are compared with respect to measurements in order to eliminate the ones that are not practical for a clinical environment. Thirdly, methods are compared in an imaging simulation environment in order to reproduce the context of noise and imaging artifacts while allowing a comparison with theoretically calculated RSP values, which will be referred to in the following as ground truth.

2.2.3.1 XCOM photon cross sections

A theoretical comparison of tissue characterization methods is performed using a set of 34 ICRU reference tissues [64]. The reference tissues with corresponding electron density are listed in table 2 of Bourque *et al.* [13] (see also corrigendum). For methods that require calibration, theoretical CT numbers of the tissue characterization phantom Gammex 467 (Sun Nuclear, Melbourne, FL, USA) are calculated and used for calibration (Hünemohr *et al.* #1 and #2, Landry *et al.* #1 and #2, Bourque *et al.*, Lalonde and Bouchard). For Han *et al.*, the calibration is done with water, polystyrene and a $CaCl_2$ aqueous solution (23%). The spectra used are from a dual source dual energy CT scanner, kindly provided by the manufacturer (Somatom Definition Flash, Siemens Sector Healthcare, Forcheim, Germany), for energies of 100 kVp and 140 kVp/Sn (Siemens custom tin filtration). Values of ρ_e and Z_{eff} (or Z_{med}) are derived for the complete set of reference tissues using the listed tissue characterization methods. Theoretical SPR values are calculated using the given electron densities and atomic compositions of the 34 human tissues. The theoretical I-values of the tissues as given from ICRP 23 [65] are calculated using the Bragg additivity rule. Although there are uncertainties in the knowledge of the I-value, such calculated theoretical SPR values provide a comparison reference to our best nowadays knowledge and form the ground truth for our study. All methods are implemented using MATLAB.

2.2.3.2 Experimental comparison with calibration phantom

A comparison based on experimental data is performed. The Gammex 467 phantom is scanned in a Siemens Somatom Definition Flash DECT scanner.

The tube voltages are 100 kV and 140 kV/Sn with tube currents 300 mAs and 232 mAs respectively. CT numbers of the tissue equivalent inserts are measured using a circular region of interest (ROI) readout (17.3 cm^3) over all slices of the phantom. The measured CT numbers are used to calibrate the methods that require calibration. Spectral knowledge is required for the spectral-based methods. The spectra of the Somatom scanner were kindly provided by the manufacturer. ρ_e and Z_{eff} are determined from the CT numbers measured in the ROIs, using each of the tissue characterization methods. A list of tissue equivalent inserts and their nominal electron densities (as specified by the phantom manufacturer) can be found in table 2 of Bourque *et al.* [13]. Again, theoretical reference values of SPRs are calculated using elemental I-values from ICRP 23 as well as the Bragg additivity rule shown in equation 2.23.

2.2.3.3 Simulated CT images

To evaluate the performance of DECT tissue characterization methods for proton therapy, it is not sufficient to test the accuracy of the methods on plastic phantoms only. Phantoms are often regular-shaped and made of similar chemical compositions, which do not entirely reproduce chemical compositions in patients. Hence, the methods need to be tested on an object resembling a patient anatomy and chemical composition of tissues, while being in a controlled environment with known reference values (here referred to as ground truth).

To simulate CT images, the software *ImaSim*, developed by Landry *et al.* [66], is used. In their previous study comparing *ImaSim* against DECT phantom images, the authors concluded that the tool is suitable to explore applications of DECT imaging in radiotherapy [67]. However, they found differences of up to 15% when comparing simulated against experimentally measured relative attenuation coefficients μ/μ_w. While discrepancies are to be expected due to the complexity of reproducing realistic CT scanners (i.e., spectra non-uniformity, reconstruction algorithms, artifact corrections, etc.), some of the features in *ImaSim* are simplified compared to the clinical reality, which could partially explain the magnitude of these differences. For the

purpose of the present study, we need to assure that *ImaSim* can reproduce most imaging artifacts encountered in clinical conditions but also CT numbers with accuracy comparable to commercial CT scanners. Therefore, a validation of the software in its ability to predict μ/μ_{w} values is necessary to assure the performance of the basic reconstruction technique and the beam hardening correction algorithm. Furthermore, since we found that the ability to reproduce realistic noise with tube current settings is questionable, a model to account for image noise is used independently from *ImaSim*.

For the image simulations, four geometries are designed. To simulate the calibration procedures, the geometry of a Gammex 467 phantom is defined, reproducing the dimensions and materials of its homogeneous disk (i.e., a diameter of 32 cm) using specifications provided by the manufacturer. A second calibration phantom is defined specifically for the method of Han *et al.*[14]. It has the same dimensions and base material as the Gammex 467 phantom, but it has only 3 inserts: water, polystyrene and $CaCl_2$ aqueous solution (23%). A third calibration phantom meant to validate *ImaSim* is defined. It has the same 13 inserts but its cylinder base is replaced by an elliptical cylinder of 32 cm width by 24 cm height. This allows us to evaluate the accuracy of the beam hardening correction in heterogeneous phantoms of irregular shapes. The fourth phantom designed has that same elliptical-shaped geometry and is a virtual patient phantom resembling a slice through the human abdomen. The virtual patient consists of various structures filled with the elemental compositions and mass densities of 15 human tissues described by Woodard and White [49, 48]. The phantom is illustrated in figure 2.1, and the list of tissues used is found in table 2.3.

All phantom scans are simulated with 3 spectra available by default in the software: 100 kVp, 120 kVp and 140 kVp/Sn. For the SECT tissue characterization techniques, the 120 kVp spectrum is used, while for DECT the 100 kVp and 140 kVp/Sn spectra are used. Image simulations are performed with infinite tube current (mAs) to disregard noise. Reconstructions are performed

with the Filtered Back Projection method using a Shepp Logan filter. For all simulations, the CT grid size is set to $0.9 \times 0.9 \times 1.0$ mm^3 voxels.

To study the impact of noise, Gaussian noise is added to simulated HU values obtained with *ImaSim*. For a consistent comparison between SECT and DECT, an equivalent amount of noise in the SECT image in terms of photon dose in water is calculated with the following relation:

$$\frac{\overline{\mu}_{\mathrm{w,SECT}}}{\Delta \mathrm{HU}^2_{\mathrm{SECT}}} = \frac{\overline{\mu}_{\mathrm{w,L}}}{\Delta \mathrm{HU}^2_{\mathrm{L}}} + \frac{\overline{\mu}_{\mathrm{w,H}}}{\Delta \mathrm{HU}^2_{\mathrm{H}}}. \tag{2.24}$$

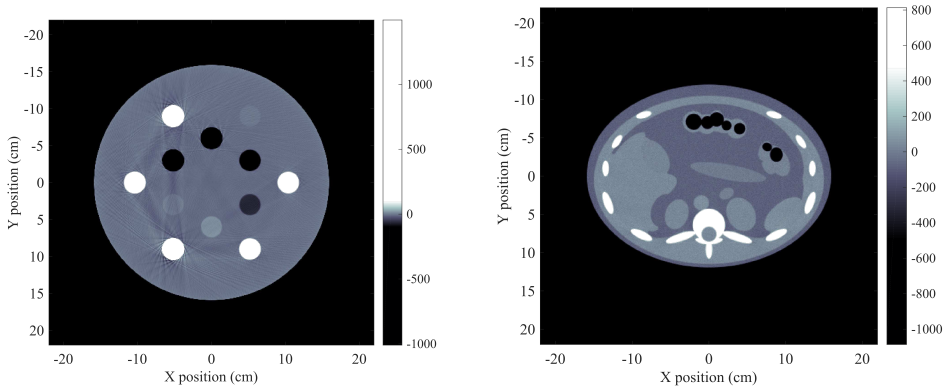
with $\Delta \mathrm{HU}_{\mathrm{SECT}}$, $\Delta \mathrm{HU}_{\mathrm{L}}$ and $\Delta \mathrm{HU}_{\mathrm{H}}$ the noise levels in SECT, DECT low kVp and DECT high kVp, respectively. The average attenuation coefficients in water $\overline{\mu}_{\mathrm{w,SECT}}$, $\overline{\mu}_{\mathrm{w,L}}$ and $\overline{\mu}_{\mathrm{w,H}}$ are calculated using the 120 kVp, 100 kVp and 140 kVp/Sn, respectively. This relation is derived using Poisson's distribution for shot noise assuming an equal dose of photons used to generate the SECT image and the DECT image pair (i.e., $D_{\mathrm{SECT}} = D_{\mathrm{L}} + D_{\mathrm{H}}$). Note that because the dose is approximately proportional to the number of photons times the mass absorption coefficient in water $\mu_{\mathrm{ab,w}}/\rho$, neglecting electron transport (hence approximating that $\overline{\mu}_{\mathrm{ab,w}}/\rho \approx \overline{\mu}_{\mathrm{w}}$) and assuming shot noise to dominate ΔHU yields equation 2.24. We study two levels of noise: 1) the low level, corresponding to SECT noise of $\Delta \mathrm{HU}_{\mathrm{SECT}} = 7$ and DECT noises of $\Delta \mathrm{HU}_{\mathrm{L}} = 12$ and $\Delta \mathrm{HU}_{\mathrm{H}} = 8$, and 2) the high level, corresponding to SECT noise of $\Delta \mathrm{HU}_{\mathrm{SECT}} = 14$ and DECT noises of $\Delta \mathrm{HU}_{\mathrm{L}} = 24$ and $\Delta \mathrm{HU}_{\mathrm{H}} = 16$.

A thorough validation of *ImaSim* is performed to assure the software to be reliable for this study. The data is validated against XCOM photon cross sections taken at the effective energies corresponding to each photon spectrum. This choice of using effective energies instead of full spectra is based on the nature of the filtered back projection reconstruction that is being used in ImaSim. The attenuation coefficients depend on the energy in each voxel. Since a spectrum is used for simulation, the energy changes along the line of response due to beam hardening. Thus, in filtered back projection, the existence of an effective attenuation coefficient is assumed and by definition different

from the average attenuation coefficient over the energy spectrum. The relative attenuation coefficients of the 13 inserts are determined and averaged over circular ROIs. For each spectrum, the effective energy E_{eff} is defined at which the residual differences between simulated and theoretically calculated relative attenuation coefficients (XCOM) are zero on average. The consistency of HU is also evaluated by comparing the simulated data as a function of the phantom shape. The averaged HU over circular ROIs of the 13 inserts are compared between the cylindrical and elliptical-shaped Gammex 467 calibration phantoms. Differences in HU are used to compare the accuracy of *ImaSim* to clinical tolerances. The tolerances were chosen to mimic clinical conditions in terms of CT reconstruction performances.

To calculate ground truth maps of SPRs in the humanoid phantom, electron densities and tissue compositions of the Woodard and White tissue database [49, 48] are used and equations 2.22 and 2.23 are applied pixelwise. For each method, tissue-specific probability distribution functions (PDFs) of SPR errors over all pixels are determined by comparing predicted SPR values to ground truth pixelwise. The PDFs are then grouped into two types of tissues: 1) soft tissues and 2) bones. This further allows determining the DECT method accuracies to predict SPRs and evaluate the effect on range uncertainties. PDFs in the absence of noise are first used to establish which DECT method is well conditioned for further comparison against SECT. The robustness to noise of the chosen DECT method is evaluated and adapted in order to determine the potential benefit of DECT over SECT in clinical conditions.

It is worth noting that the ground truth SPR map of the virtual humanoid phantom is not affected by noise or imaging artifacts. However, limitations caused by the CT grid size are left aside by avoiding analyzing data adjacent to interfaces, this way assuring voxels to be homogeneous. Also, because the accuracy of reference values is limited by the Bragg additivity rule, the present study leaves aside uncertainties related to the I-value by (directly or indirectly) using the same rule to predict SPR. This way, the present work focuses mainly



(a) **(b)**

Figure 2.1: Geometries used for the *ImaSim* simulation: (a) a simulated 140 kVp/Sn CT image of the calibration phantom (resembling the Gammex RMI 467) with added noise (1σ = 16 HU), showing artifacts reproduced by *ImaSim*; (b) a simulated 100 kVp/Sn CT image of the virtual humanoid phantom geometry.

on the effect of CT noise and artifacts, leaving the effects of CT grid size and I-value aside.

Table 2.3: List of the 15 human tissues specified by Woodard and White [49, 48] used to simulate CT images and calculate ground truth SPRs. The I-values are calculated using the Bragg additivity rule from the elemental composition using equation 2.23 and I-values recommended by ICRP [65]

Tissue number	Tissue name	Electron density relative to water	I-value (eV)
1	Adipose tissue	0.951	64.780
2	Adrenal gland	1.025	70.835
3	Aorta	1.038	75.160
4	Blood, whole	1.050	75.203
5	Gallbladder bile	1.026	75.245
6	Kidney	1.040	74.286
7	Liver	1.041	74.355
8	Mammary gland	1.014	70.294
9	Muscle, skeletal	1.040	74.621
10	Ribs 6th and 2nd	1.347	90.722
11	Small intestine wall	1.024	74.285
12	Spleen	1.051	74.980
13	Stomach	1.042	74.194
14	Vertebral column C4	1.355	91.218
15	White matter	1.034	73.126

2.2.4 Evaluation of range uncertainties

The impact of the DECT methods on proton beam range uncertainty is evaluated using numerical models. To evaluate the impact in soft tissues, a water equivalent path length (WEPL) based method [68] is used in combination to SPR error sampling at depth increments of 1 mm, to be consistent with the largest dimension of CT voxels used in the *ImaSim* simulations. For each tissue characterization technique (SECT or DECT), beam range errors are sampled repeatedly by individually sampling SPR errors at each depth increment of 1 mm with PDFs determined from results of the simulated CT images in soft tissues. This way, the performance of the method in extracting SPR from simulated CT images determines the probability distribution of SPR errors. Each statistical sample of range error is calculated analytically from a random array of SPR errors through which the beam is transported. For a given beam energy, depth-dose curves of pristine proton beams are calculated by remapping the depth-dose curve in water, initially calculated with the PSTAR lookup table [69], to the array of WEPL values associated to the random array of SPR error values set in each 1 mm depth increment. That is, one range error sample corresponds to one array of SPR errors set in each depth increment, as illustrated in fig 2.2. The calculated range is then compared to the expected range in water (i.e., without SPR errors) to estimate the range error for that random array of SPR errors. In the dose falloff, the final depth increment is reduced to the proton track-end in order for the result not to be influenced by the size of the CT grid. The statistical distributions of beam range errors in soft tissues are used to estimate the 95% confidence intervals for each tissue characterization technique. The same rationale is used to evaluate the impact of SPR errors in bones and its effect on the range uncertainty. The error in range caused specifically by transport in bones is attributed to its uncertainty in energy loss through them. For a given bone thickness, a number of depth increments are defined (and again set to 1 mm) and a random array of SPR errors is sampled with the PDFs determined in bones from the simulated CT

images. Energy loss errors in bones are estimated with Bethe's formula (equation 2.22). The calculated errors on energy loss are translated into range shift by using the PSTAR energy-range lookup table in water[69] as a function of the beam energy.

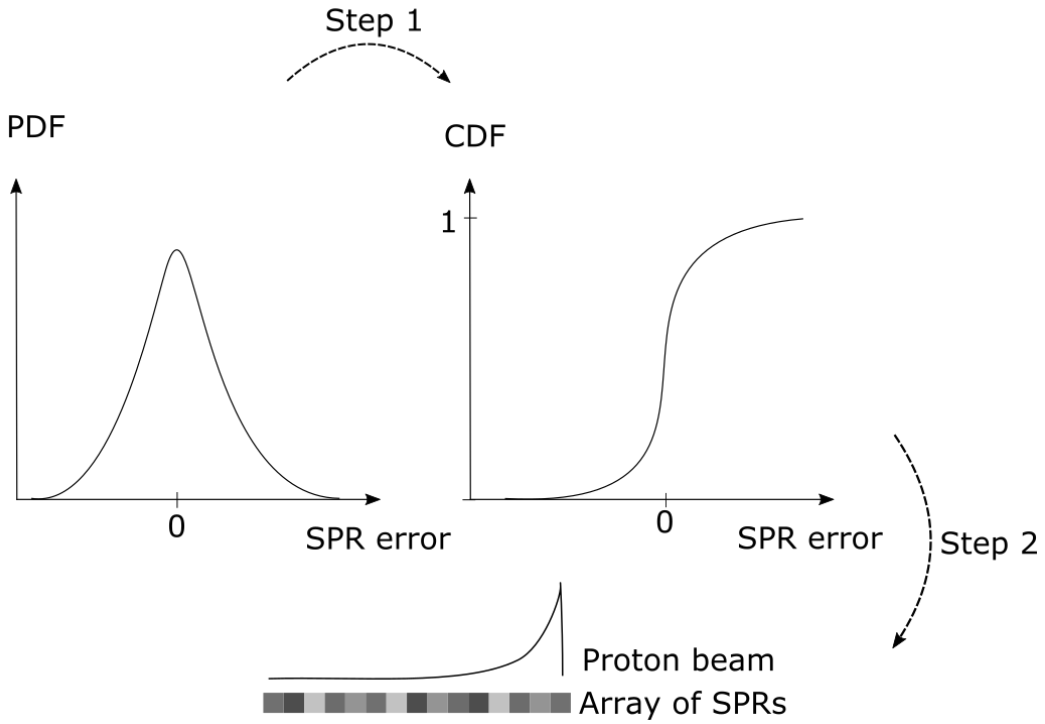


Figure 2.2: Schematic illustration of the range uncertainty determination. The PDFs of SPR errors are converted into cumulative density functions (CDFs). From the CDFs, SPR errors are randomly sampled and converted into an array of SPRs. A proton beam with a fixed energy is transported through this array. The range of this beam is compared to the nominal range of the beam and the difference is recorded. This process is repeated 2200 times to gain a mean range error and standard deviation.

2.3 Results

2.3.1 Theoretical comparison of tissue characterization methods

All methods are applied to theoretical attenuation coefficients to predict the SPR of 34 human reference tissues. The residual analysis between predicted and theoretical SPR values is found in table 2.4. All methods are capable of

predicting the SPR of human tissues within 1% under ideal conditions. The methods by Bazalova *et al.* and Bourque *et al.* appear to give the most accurate SPR predictions within a theoretical setup, this considering the negligible bias (i.e., the mean error) and the smallest root mean square error, although Bourque *et al.* contains its errors within the smallest unbiased interval (i.e., less than ±0.4%). The method by van Abbema *et al.* introduces a bias to tissues with a high EAN. In their publication, van Abbema *et al.* discovered that EANs determined from their method suffer a systematic deviation. Therefore, the authors suggest that their method should only be used for electron density determination.

Table 2.4: Statistics of residual errors of theoretically determined SPRs for 34 human reference tissues using the investigated formalisms.

Method	Min (%)	Max (%)	Mean (%)	RMS (%)
Bazalova *et al.*	-0.47	0.26	-0.02	0.16
Landry *et al.* #1	-0.46	0.33	-0.06	0.20
Landry *et al.* #2	-0.72	0.34	0.17	0.17
Hünemohr *et al.* #1	-0.46	0.33	0.03	0.19
Hünemohr *et al.* #2	-0.43	0.29	0.16	0.16
Bourque *et al.*	-0.38	0.38	0.04	0.16
Van Abbema *et al.*	-0.84	-0.04	-0.30	0.41
Han *et al.*	-0.55	0.60	0.01	0.23
Lalonde and Bouchard	-0.48	0.54	-0.01	0.19

2.3.2 Experimental comparison of tissue characterization methods

All methods are used with scanned images to predict the SPR of the Gammex 467 phantom. The results, displayed in table 2.5, are compared to theoretically calculated SPR values for the 13 inserts. The spectral based methods (Bazalova *et al.*, van Abbema *et al.*) suffer from a systematic bias in the region of higher EAN. This problem was addressed by both authors. Bazalova *et al.* suggested a semi-empirical correction to the subset of data points that are affected by this bias. Van Abbema *et al.* suggest an LWF for every pixel in the image. Although this LWF is applied here, we still observe a bias for higher-Z materials, which was discussed in the paper by van Abbema *et al.*

and is addressed above. During our study, we found that the calculation of the LWF and the process numerically solving $\mu_{\mathrm{L}}/\mu_{\mathrm{H}}$ requires high computational effort and time.

Calibration-based methods show a good overall performance in a phantom setup. The methods by Landry *et al.* #1 and #2, Hünemohr *et al.* #1 and Bourque *et al.* describe SPRs of phantom materials within ±2%. This residual analysis compares both approaches (spectral- and calibration-based) and is intended to show that spectral-based methods need further consideration to reach the accuracy of calibration based-methods. Despite that both approaches can reach similar theoretical performances (see table IV), the calibration-based methods yield more accurate residuals with experimental data (see table V), since the spectral information is likely not to be representative of the actual spectrum. Also, due to beam hardening effects, the spectrum is not unique in space for all projections. Therefore, one could assume the existence of an effective spectrum giving optimal experimental results. Fitting the spectrum to the experiments would improve the model, but would end up being considered as a calibration-based method. The observed discrepancies between theoretically calculated SPRs (i.e., based on electron densities and compositions provided by the vendor) and those found using the calibrations have three major uncertainty components: 1) experimental uncertainties, 2) uncertainties in the phantom composition and 3) uncertainties in the models themselves. With the residual analysis performed herein, we compare the uncertainties of the models consistently without changing the other first two sources of uncertainties, therefore consistently comparing the models under the same conditions. It is worth noting that the method of Lalonde and Bouchard is designed to describe human tissues only, as the principal components used in the material decomposition are not applicable to the Gammex phantom materials. This might explain some of the large differences reported in table 2.5, although the method is overall unbiased with a negligible mean error. Also, note that the method of Han *et al.* and is not included in the experimental comparison as

the technique requires the use of $CaCl_2$ solutions which was not considered in the present study.

Table 2.5: Statistics of residual errors of experimentally determined SPRs of the Gammex 467 calibration phantom using the investigated formalisms.

Method	Min (%)	Max (%)	Mean (%)	RMS (%)
Bazalova *et al.*	-1.49	4.29	0.57	1.67
Landry *et al.* #1	-1.61	1.78	-0.11	0.80
Landry *et al.* #2	-1.52	1.20	-0.12	0.70
Hünemohr *et al.* #1	-1.73	1.25	-0.23	0.81
Hünemohr *et al.* #2	-2.22	1.93	-0.10	1.21
Bourque *et al.*	-1.57	1.12	-0.25	0.68
Van Abbema *et al.*	-2.04	8.55	1.12	3.19
Han *et al.*	-	-	-	-
Lalonde and Bouchard	-2.52	2.82	0.06	1.66

2.3.3 Comparison of tissue characterization methods based on simulated CT images

2.3.3.1 Validation of *ImaSim*

The ability of *ImaSim* to reproduce attenuation coefficients is evaluated on the results obtained with the cylindrical and elliptical-shaped calibration phantoms. In comparing results of the cylindrical phantom with XCOM cross sections data, the worse case scenario is found for the 100 kVp spectrum ($E_{\text{eff}} = 69.3\,\text{keV}$) with errors ranging from -0.9% to 1.1%, and a root mean square error of 0.7%. The same analysis with experimental data of the Gammex 467 phantom scanned with a Siemens Somatom Flash Definition dual-source CT yields mean absolute errors of range from -1.7% to 1.9%, and a root mean square error of 1.0%, for the 100 kVp spectrum ($E_{\text{eff}} = 71.6\,\text{keV}$). Because experimental data are expected to be higher than numerical simulations due to additional sources of uncertainties, this shows that *ImaSim* is reliable for cylindrical geometries. In its performance with the elliptical-shaped calibration phantom, the worst discrepancies on average HU values between the cylindrical and elliptical-shaped phantoms are found to be for the 100 kVp spectrum and range between -2.2 and 0.5 HU as well as 7.1 and 37.5 HU

for the plastics equivalent to soft tissues and bones, respectively. However, because only two bones are defined in the virtual humanoid phantom, i.e., vertebral column and ribs, two of the materials in the calibration phantom are out of range in terms of density. Removing these in the analysis yields a maximum discrepancy of 10.4 HU. These results show that the beam hardening correction is acceptable for soft tissues, compared to typical vendor recommendation of ±4 HU for water. However, errors are slightly higher in bones than expected. But when comparing the elliptical-shaped results against XCOM cross sections with the same effective energy as found for the cylindrical phantom ($E_{\text{eff}} = 69.3\,\text{keV}$), leaving the two high-density inserts aside (i.e., SB3 and CB2 - 50%) yields errors ranging from -1.7% to 1.3% with a root mean square error of 1.0%. This is comparable to experimental results obtained with the cylindrical Gammex 467 phantom. Therefore, we conclude that *ImaSim* is an acceptable tool for the present study.

2.3.3.2 Estimated probability distribution functions of SPR errors

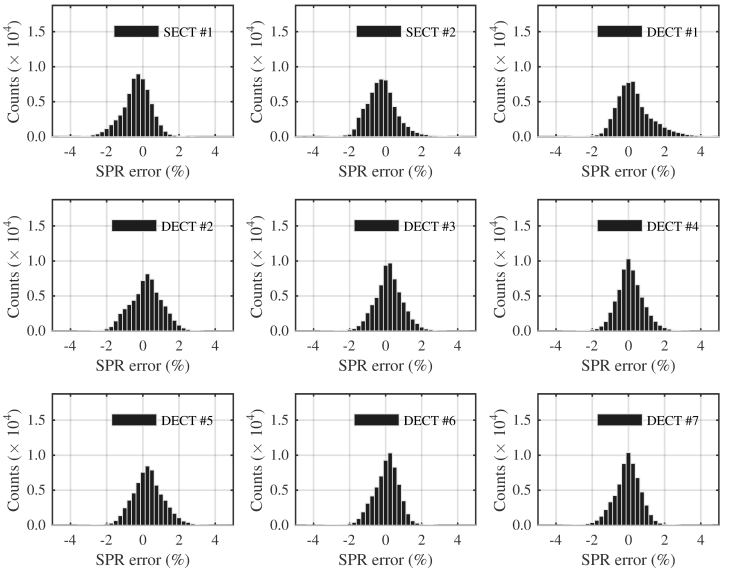
To reproduce clinical use, only calibration-based methods are used to predict SPRs from simulated DECT images pixelwise. The differences between predicted SPR maps and ground truth SPR values are analyzed. The SECT method proposed by Schneider *et al.* (1996) serves as a gold standard for ρ_e-Z formalisms. PDFs of SPR errors in the absence of noise are displayed in figure 2.3. The statistics of the methods is summarized in table 2.6. For soft tissues, all investigated DECT methods predict SPRs with a smaller mean error than the SECT method of Schneider *et al.* (1996), therefore introducing a smaller bias and decreased errors on proton range. Among our implementations, the method by Bourque *et al.* is found to have the smallest mean error, thus introducing a quasi-null bias on proton range prediction, as well as the smallest standard deviation on SPR. For bones, not all DECT methods have a smaller mean error than SECT methods. Three DECT methods introduce a higher bias than the method of Schneider *et al.* (1996), and four introduce a higher

bias than the SECT method of Schneider *et al.* (2000), which is shown to improve the characterization of bones compared to the gold standard SECT. The method of Lalonde and Bouchard was found to have a quasi-null bias and the smallest standard deviation. To determine if the population means of the probability density functions are statistically different, we performed pair-wise Welsh's t-tests. In soft tissues, for each pair of PDFs, we found p-values smaller than 10^{-5}, indicating that all distributions are significantly different from each other ($p < 0.05$), with one exception. The distributions derived from Hünemohr *et al.* #1 ($\mu = 0.1068$) and Han *et al.* ($\mu = 0.1148$) are statistically similar ($p = 0.293$). For bones, we found that all distributions are significantly different from each other ($p < 0.05$), with the exception of Hüemohr *et al.* #1 and Landry *et al.* #1 ($p = 0.064$). The Welsh's t-test is designed for normal distributions. While the PDFs for soft tissues can be assumed almost normally distributed, this is not the case for the PDFs for bone, which should be considered when interpreting the results of the statistical significance.

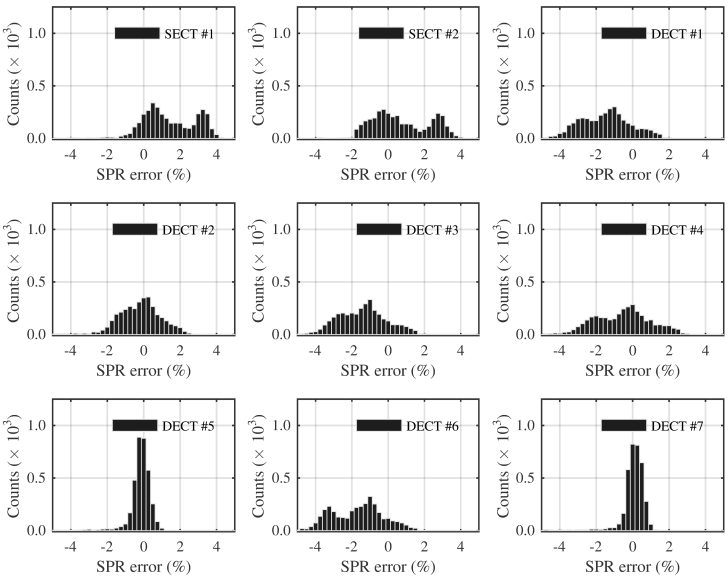
It is worth noting that the methods suitable to predict MC inputs (i.e., Landry *et al.* #2, Hünemohr *et al.* #2 and Lalonde and Bouchard) do not perform better in soft tissues than the $\rho_e - Z$ decomposition method of Bourque *et al.*. These results lead to believe that the intermediate step of assigning elemental weight fractions before calculating SPR might not be optimal as it can reduce the accuracy of the estimation. However, the potential improvement in dose calculation using MC simulation over analytic tools used commercially are not shown explicitly in these results. Therefore the DECT techniques suitable for MC should not be literally compared with the ones suitable for analytic methods.

2.3.4 Proton beam range error estimations

Results are calculated for each noise level separately, i.e., none, low and high. It is worth noting that for accurate estimations of range error confidence intervals, a sufficiently large number of samples is required to get a smooth behavior of the results as a function of the beam energy and/or bone thickness. The



(a)



(b)

Figure 2.3: PDFs of SPR errors generated with each method applied on the simulated images in the absence of noise: (a) soft tissues and (b) bones. The SECT methods #1 and #2 are Schneider *et al.* (1996) and Schneider *et al.* (2000), respectively. The DECT methods from #1 to #7 are Landry *et al.* #1, Han *et al.*, Hünemohr *et al.* #1, Bourque *et al.*, Hünemohr *et al.* #2, Landry *et al.* #2 and Lalonde and Bouchard, respectively. The display of errors is reduced to within ±5%, although larger errors occur.

Table 2.6: Statistics of the PDFs of SPR errors of all investigated tissue characterization methods in the absence of noise: the mean ($\hat{\mu}$) and the standard deviation ($\hat{\sigma}$).

Method	Soft tissues		Bones	
	$\hat{\mu}$ (%)	$\hat{\sigma}$ (%)	$\hat{\mu}$ (%)	$\hat{\sigma}$ (%)
Schneider *et al.* (1996)	-0.43	1.42	1.34	1.61
Schneider *et al.* (2000)	-0.29	1.49	0.65	1.72
Landry *et al.* #1	0.27	1.40	-1.72	1.87
Han *et al.*	0.11	1.34	-0.41	1.71
Hünemohr *et al.* #1	0.11	1.28	-1.64	1.83
Bourque *et al.*	0.02	1.25	-0.77	1.95
Hünemohr *et al.* #2	0.23	1.26	-0.34	1.22
Landry *et al.* #2	-0.04	1.27	-1.92	1.89
Lalonde and Bouchard	-0.13	1.27	-0.08	1.14

number of samples per method and per beam energy is set to N=2200, totaling 415 800 range error sampling for soft tissues for all 3 levels of noise. For bones, the number of samples is 2 079 000 since five bone thicknesses are investigated, totaling about 2.5 millions of range error samples.

2.3.4.1 Comparison of DECT methods in the absence of noise

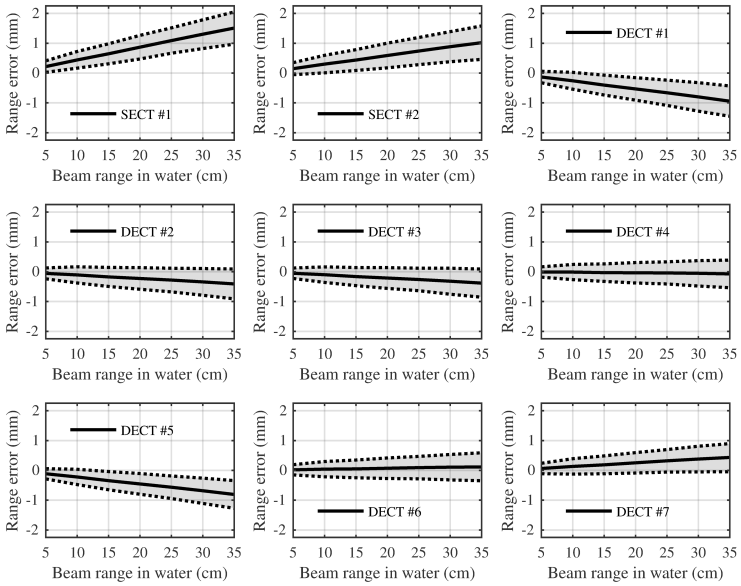
Two independent sources of range uncertainties are evaluated from PDFs. The first effect is the range error limited by the precision of SPR predictions in soft tissues. The second effect in the range error caused by proton beam transport through bones before being aimed at a tumor (located in soft tissue). Resulting effects on range errors are shown in figure 2.4. The effects are consistent with the statistics of the PDFs reported in table 2.6. In soft tissues, both SECT methods are systematically biased, while most of the implementations of the DECT methods show smaller bias and similar or smaller 95% range error distribution, with five out of seven methods having low bias: Han *et al.*, Hünemohr *et al.* #1, Bourque *et al.*, Landry *et al.* #2 and Lalonde and Bouchard. The smallest range errors were found in the method by Bourque *et al.*, with maximal beam range errors within -0.54 mm and 0.39 mm, with a probability of 95%, for beam energies corresponding to ranges in water of up to 35 cm. For the impact of transporting proton beams through bones,

both SECT methods are systematically biased, while in our implementations four of out seven DECT methods yield low bias: Han *et al.*, Bourque *et al.*, Hünemohr *et al.* #2 and Lalonde and Bouchard. The smallest range errors were found in the method by Lalonde and Bouchard. It shows maximal beam range errors within -0.91 mm and 1.05 mm (with a probability of 95%) for bone thicknesses up to 5 cm and for beam energies corresponding to a range in water of up to 35 cm.

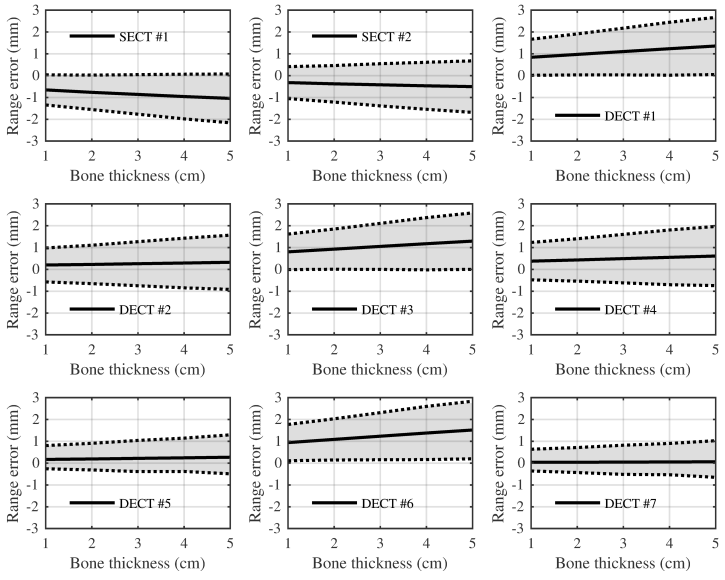
2.3.4.2 Range uncertainties in clinical conditions: the impact of noise

The impact of noise on beam range uncertainties is evaluated by applying the range error estimator models on PDFs calculated with two levels of noise. For soft tissues, the SECT method used is the gold standard method of Schneider *et al.* 1996 and the DECT method is the one of Bourque *et al.*, but adapting its fit parameter of the dual-energy index (DEI) versus Z to lower order to make it more robust to noise (i.e., $K = 3$ instead of $K = 5$). When noise is present in the image, the values for the DEI can fall out of the calibration domain. By choosing a lower fit order, we are able to control the behavior of the calibration curve outside the calibration domain. An alternative approach would be to use the high fit order ($K = 5$) for values within the calibration domain and additionally describe values outside the calibration domain using a linear extrapolation. For bones, the SECT method used is Schneider *et al.* 1996 and the DECT method used is the method of Lalonde and Bouchard without any modification. The selection of DECT methods to investigate the impact of noise was based on their performance without noise, showing the smallest mean error in soft tissues and bones, respectively. Since the theoretical robustness of methods is investigated, all here presented comparisons are made under idealistic conditions. Therefore, selecting the method with the smallest differences is a suitable criterion.

Results are shown in figure 2.5. The mean errors and boundary values of the 95% confidence interval of range errors in SECT and DECT are compared.



(a)



(b)

Figure 2.4: Comparison of estimated range errors in tissues from CT data excluding noise for: a) soft tissues, and b) bones. The plain line shows the mean error values and the dotted lines show the boundaries of the 95% confidence intervals of range errors. The energy used the effect in bones is 196 MeV, corresponding to a range in water of 25 cm. The method's numbering is the same as in figure 2.3.

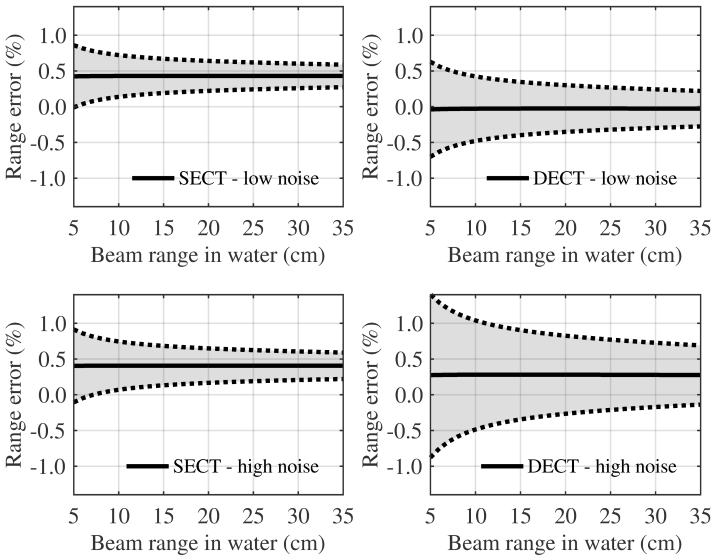
For soft tissues, results are displayed as a function of the beam energy, reported in terms of range in water. The comparison shows that for the low level of CT noise, range errors with DECT methods are unbiased compared to SECT, with

interval boundary values closer to zero. For the high level of CT noise, the DECT interval is slightly biased and the boundary values are much higher than for SECT, which sensitivity to noise is small. In the absence of noise, maximum range absolute errors with DECT are decreased by about 0.5% relative to the beam range in water, while for the low level of CT noise they are reduced by up to 0.4% relative to the beam range in water. However, for the high level of noise SECT had smaller range uncertainties than DECT, despite its bias in predicting the range.

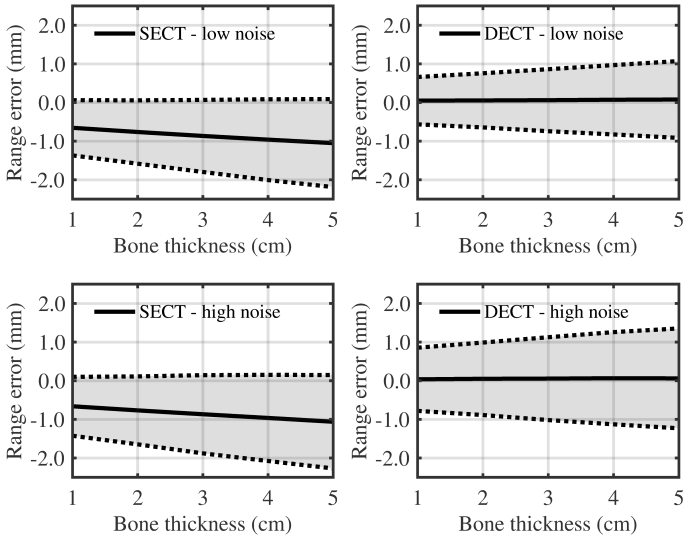
For bones, results are displayed as a function of bone thicknesses through which a 196 MeV beam is transported. The comparison shows that for the low and high levels of noise, DECT errors are unbiased compared the SECT, with interval boundary values closer to zero. Between 1 and 5 cm bone thickness, maximum range absolute errors are reduced by values of up to about 0.6 to 1.1 mm with DECT. For the high level of noise, the same calculations (not shown here) lead maximum range absolute errors reductions between 0.5 and 0.9 mm for bone thicknesses between 1 and 5 cm, respectively.

2.4 Discussion and conclusion

In the present study, the potential of DECT is evaluated over SECT in the context of proton beam range prediction. Nine different techniques are compared in their ability to predict proton SPRs. The methods are implemented and evaluated in three different contexts to evaluate their theoretical foundation (i.e., with XCOM cross sections data), their practicality in a clinical environment (i.e., with measurements) and their performance with a patient-like geometry under constraints of CT artifacts and noise (i.e., *ImaSim* simulations and Gaussian noise model). The first two contexts allow reducing the number of suitable methods to seven. The performance of the DECT methods with a humanoid phantom is first estimated in the absence of noise to allow choosing techniques being the most robust to CT artifacts, i.e., Bourque *et al.* for soft tissues and Lalonde and Bouchard for bones. It is worth noting



(a)



(b)

Figure 2.5: Comparison of range error statistics between SECT and DECT for low and high CT noise in a) soft tissues, and b) bones. The dotted lines represent the boundary values of the 95% confidence interval and the plain line is the average error.

that all methods are implemented to the best of our knowledge, based on the publications available in literature. We use the theoretical comparison based on XCOM data, as well as the experimental data, as an indicator to assure that the methods are implemented properly. Our results reproduce values that

were quoted by the authors of each method, leading to the conclusion that all methods should be correctly implemented.

The most clinically-relevant results of this study are the ones where CT artifacts and noise are present. Range error estimations clearly demonstrate the advantages of DECT over SECT in the presence of low CT noise, since SECT is generally more robust to noise due to the mathematical nature of its techniques (i.e., linear models). Overall, one could expect DECT to reduce range uncertainties (to the 95% confidence level) by about 0.4% in soft tissues, and up to about 1 mm for beams of therapeutic energies transported through bones. For high levels of CT noise, the benefits of DECT can be lost over the robustness of SECT in soft tissues. While this is expected due to the mathematical complexity of DECT techniques, it is yet to be demonstrated that some techniques could be further adapted for high CT noise. For instance, Bourque *et al.* is used in soft tissues with minimal adaptation (i.e, just changing $K = 5$ to $K = 3$ in the DEI conversion to Z), and this could explain why it is only robust to low noise levels. As for the method of Lalonde and Bouchard, it is surprising that despite no adaptation it stills outperforms the SECT gold standard for low or high noise levels. This could suggest that an effort in adapting the method for the presence of noise could yield even better results. The results suggest that DECT-predicted SPR can benefit from an increase in mAs defined in the scanning protocol. Therefore we recommend investigating SPR uncertainties before establishing a clinical DECT protocol for radiotherapy planning. We would like to emphasize that errors arising from spectral differences between the calibration and patient scan are not taken into account here. Therefore we recommend performing the calibration for each scanner model and scanning protocol individually.

While the benefits of DECT over SECT are expected to be improved by refined robustness to noise, one could also seek for more sensible values in range uncertainties to be obtained with a more realistic dose calculation model, such as MC simulations. However, performing such a study with MC transport

simulations is rather difficult, yet impossible, as a high number of range error samples is required (i.e., nearly 2.5 million in this study), which in the context of cross sections become multidimensional rather than simply the SPR error, requiring to redefine a set of materials and a full calculation (with millions of histories) for each sample. Nonetheless, it is quite conceivable that the numbers estimated in the present study are realistic due to the consistency of the methods. The simulation of CT images using *ImaSim* has the advantage of allowing SPR estimation with various techniques in a controlled and consistent environment, with focus on CT artifacts and noise, leaving the effects of CT grid size, uncertainties in I-values and other sources aside. Finally, while the WEPL-based model is not entirely accurate, it is still used consistently and therefore should yield correct estimations of errors.

A simplified interpretation of the results presented in the present study allows comparison with the topical review by Paganetti [9]. In that publication, uncertainties in CT conversion to tissue as well as CT imaging and calibration each contribute to 0.5% of the range uncertainty (1.5σ), and the overall uncertainty recommended for proton beam range is 2.7% + 1.2 mm. While adding the two uncertainty sources in quadrature yields about 0.9% for a significance level of 95%, this value corresponds to the maximum error found in the present study at the highest noise level for SECT. From this, we could conclude that the recommended uncertainty with DECT should be reduced to 2.4% + 1.2 mm (i.e., reporting the 95% level of confidence, corresponding to a statistical significance of 1.5σ). But a closer look at the present results suggests a deeper analysis, which is addressed in figure 2.6. Here we illustrate the main advantage of unbiased range errors, as it allows reducing the size of the margins. Indeed, DECT has the advantage of reducing uncertainties as only the interval boundary needs to be considered as an uncertainty for each direction with respect to the beam, conversely to using the maximum absolute error in SECT.

The method proposed in the present study provides a more detailed es-

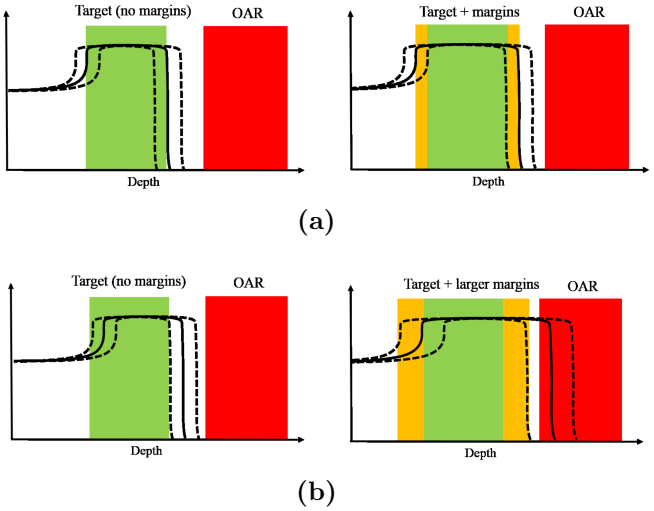


Figure 2.6: Illustration of the effect of range uncertainties on the definition of margins adjacent to still targets in two situations: a) unbiased range uncertainties, and b) biased range uncertainties. In each figure, the left graph shows the spread-out Bragg peak (SOBP) adapted for the target, while in the second the SOBP it is adapted to the target plus margins, accounting for range uncertainties. In this example, the systematic bias of the error doubles the size of the margins and compromises OAR sparing.

timation of range uncertainties than more simplistic rules used in the clinic (i.e., set to 3.5% of the range in water for all energies). An interesting result in figure 2.5a shows that range uncertainties relative to the beam range in water are larger for smaller energies. This can be explained by the fact that the smaller the energy, the smaller the number of voxels contribute to the average SPR. This way, the uncertainty on the average SPR is inversely proportional to the square root of the number of voxels traversed. And because the range relative to that of water equals the inverse of the average SPR, with a few manipulations we show that the relative range uncertainty is given by [13]

$$\frac{\Delta R}{R} = \frac{\Delta \mathrm{SPR}_{\mathrm{ave}}}{\mathrm{SPR}_{\mathrm{ave}}} = \sqrt{\frac{\Delta x}{R}} \frac{\Delta \mathrm{SPR}}{\mathrm{SPR}_{\mathrm{ave}}}, \tag{2.25}$$

with Δx the size of the voxels in which SPR values are assumed homogeneous and ΔSPR the uncertainty on SPR in each voxel. This relation predicts that for a fixed CT grid size and uncertainty on SPR the relative range uncertainty in soft tissue (where $\mathrm{SPR}_{\mathrm{ave}}$ is approximately constant) is inversely proportional

to the square root of the range, which is consistent with results shown in figure 2.5a.

Finally, although it could be possible to improve SECT methods notably by using Schneider *et al.* 2000 or attempting to correct for the bias, the present study suggests that DECT can go beyond the capabilities of SECT in the context of proton therapy. However, noise remains a major limiting factor and needs to be carefully addressed if the patient imaging dose is to be kept to the same level as in conventional radiotherapy treatment planning. We conclude that DECT has substantial potential for reducing range uncertainties in proton therapy and that further developments of DECT methods should focus on their robustness to noise since mathematical formalisms might have found their full maturity at the present time. Also, it is expected that DECT methods based on raw-data should enable the reduction of CT artifacts, and therefore range uncertainties. Moreover, improvements in CT grid size (i.e., such in future developments in spectral CT) could help to improve the precision of proton therapy planning.

Before proton therapy planning with DECT predicted SPR values is possible, a thorough experimental validation of any applied method should be performed. If SPR values are used directly as input for treatment planning, the work presented in this chapter suggests the use of the method by Bourque *et al.* [13]. It was shown that the method has a low biased and can be adapted to be robust to noise. In the case of MC planning where elemental compositions are desired as planning input, this work suggests the use of the Lalonde and Bouchard [26] method due to its low bias in predicting soft tissue and bone SPR and ranges. For this method, further work is performed to ensure robustness to image noise.

Chapter 3

Experimental validation of two dual-energy CT methods for proton therapy using heterogeneous tissue samples

The work presented in this chapter can be found in the following journal articles:

1. Zang, R., **Bär, E.**, Jee, K. W., Sharp, G. C., Flanz, J., Lu, H.-M. (2017) Investigation of real tissue water equivalent path lengths using an efficient dose extinction method. *Phys. Med. Biol.* 62: 5640-5651. doi: 10.1088/1361-6560/aa782c.

2. **Bär, E.**, Lalonde, A., Zhang, R., Jee, K. W., Yang, K. , Sharp, G. C., Liu, B., Royle, G., Bouchard, H., Lu, H. M. (2018) Experimental validation of two dual-energy CT methods for proton therapy using heterogeneous tissue samples. *Medical Physics* 45: 48-59. doi: 10.1002/mp.12666.

It was presented at the following meetings:

3. **Bär, E.**, Lalonde, A., Zhang, R., Jee, K. W., Yang, K. , Sharp, G. C., Liu, B., Royle, G., Bouchard, H., Lu, H. M. (2016) Experimental

validation of the DECT stoichiometric calibration for proton treatment planning using real tissue samples. Boston, *AAPM New England Chapter, Peter Neurath Young Investigator Symposium.*

4. **Bär, E.**, Lalonde, A., Zhang, R., Jee, K. W., Yang, K. , Sharp, G. C., Liu, B., Royle, G., Bouchard, H., Lu, H. M. (2016) The Impact of Using Dual-Energy CT for Determining Proton Stopping Powers: Comparison Between Theory and Experiments. Washington D.C., *58th Annual Meeting of the AAPM. Medical Physics* 43 (6Part35): 3756-3756. doi: 10.1118/1.4957542.

5. Zang, R., **Bär, E.**, Jee, K. W., Sharp, G. C., Flanz, J., Lu, H.-M. (2016) Efficient dose extinction method for water equivalent path length (WEPL) of real tissue samples for validation of CT HU to stopping power conversion. Washington D.C., *58th Annual Meeting of the AAPM. Medical Physics* 43 (6Part11): 3452-3452. doi: 10.1118/1.4956101.

6. **Bär, E.** Lalonde, A., Zhang, R., Jee, K. W., Yang, K. , Sharp, G. C., Liu, B., Royle, G., Bouchard, H., Lu, H. M. (2016). The Impact of Using Dual-Energy CT for Determining Proton Stopping Powers of Real Tissues. London, *3rd PPRIG workshop.*

7. **Bär, E.**, Lalonde, A., Royle, G., Lu, H. M., Bouchard, H. (2017) On the performance of dual-energy CT for determining stopping power ratios for proton radiotherapy. Montreal, *9th QBIN Scientific Day.*

8. **Bär, E.**, Lalonde, A., Royle, G., Lu, H. M., Bouchard, H. (2018) Experimental quantification of range uncertainties arising from lateral inhomogeneities. Nashville, *60th Annual Meeting of the AAPM, presentation as electronic poster.*

Contribution of Authors: Arthur Lalonde developed the codes for the DECT method referred to in the following as Bayesian ETD (Lalonde *et al.*, 2017). He also performed the MC simulations for this study. Rongxiao Zhang,

Kyung-Wook Jee and Gregory Sharp were involved in the development and initial testing of the measurement methods applied in this study. Kai Yang and Bob Liu provided access to the CT scanner and were involved in the acquisition of DECT images. Gary Royle provided input on the impact of this work and initiated the collaboration with MGH, making this project possible. Hugo Bouchard provided supervision and valuable input on every aspect of data processing and analysis. Hsiao-Ming Lu supervised the experimental aspect of the project and provided guidance during data collection. I performed the main part of the data collection, this included the design and performance of experiments, developing codes and tools for data processing and the final analysis.

3.1 Introduction

In this chapter, two main contributions of range uncertainties are investigated. The main part of the chapter focuses on the uncertainties related to CT imaging and conversion to tissue, and the question of how DECT can improve range predictions compared to SECT. A smaller part focuses on the uncertainties introduced by the range calculation algorithm. Here, a simple ray tracing approach to estimate the range from DECT data is compared to a more sophisticated MC method.

3.1.1 Range uncertainties from CT imaging and conversion to tissue

In clinical practice, the conversion from CT numbers (in HU) to SPRs for SECT is often performed using the stoichiometric calibration method proposed by Schneider *et al.* [10]. The CT scanner is calibrated for each specific X-ray spectrum and CT scan protocol by acquiring the CT numbers of a calibration phantom composed of plastic materials of known composition and density. The CT-number-to-SPR calibration curve is established for each CT scan protocol by applying the calibrated model on a set of human reference

tissues [49]. As stated in section 2.1, this calibration method was validated experimentally by Schaffner and Pedroni [50]. They performed a SECT scan of different biological samples (liver, muscle, spleen, heart, brain, adipose, kidney, blood) and applied the CT-number-to-SPR conversion to the samples. From the resulting SPR maps, they calculated the WER of a proton beam traveling through the samples and compared it to measured values of the WER. This study reports range errors caused by CT scanning and calibration of 1.1% in soft tissues and 1.8% in bones, taking beam hardening into account (expected 1% without beam hardening artifacts).

DECT was proposed in literature to potentially increase the accuracy of SPR predictions. Yang *et al.* [23] performed a theoretical study comparing SECT and DECT determined SPRs of standard human biological tissues. They proved the theoretical superiority of DECT over SECT, especially when the electron density and elemental compositions of the investigated tissue vary from reference tissues.

Several formalisms for converting CT numbers from DECT to SPR were proposed in literature. Most of these formalisms are parameter-based and aim at the extraction of ED and EAN [52, 53, 15, 16, 13, 21, 18] or photon absorption cross section [17], which are used to estimate the mean excitation energy (I-value) or the stopping number, respectively. Alternative methods parametrize the I-value [14] or the SPR [19] directly, while other groups focus on calibrating pseudo-monoenergetic images derived from DECT to predict the SPR of tissues [22]. Another class of formalisms aims at extracting elemental compositions as inputs for MC simulations [25, 24, 26, 47]. In chapter 2, we performed a theoretical study on the potential of DECT to reduce proton range uncertainties [70]. For a virtual phantom filled with human reference tissues, we found that DECT can reduce beam range uncertainties by about 0.4% in soft tissues and up to 1 mm (as reported for a 5 cm thick slab) for

therapeutic energies in bones. Additional studies [71, 72] performed treatment planning on phantom and patient data sets, comparing treatment plans calculated using SECT-derived and DECT-derived SPRs. Zhu and Penfold [72] demonstrate a dose difference of up to 8% between a SECT-based and a DECT-based plan. Hudobivnik *et al.* [71] evaluated range differences between SECT-based and DECT-based plans for five head trauma patients, concluding group median relative range differences of -1.4%.

The presented study aims at validating DECT for estimating SPR values in biological tissues. We utilize animal tissue samples consisting of single organs and a variety of different animal bones to measure the WER of the samples in a proton beam using the dose extinction method proposed by Zhang *et al.* [73]. The measured WER is then compared to a calculated WER predicted from SECT and two different DECT formalisms. These formalisms include one ED-EAN method (Bourque *et al.*) and one eigentissue decomposition (ETD) method (Lalonde *et al.*).

3.1.2 Range uncertainties from the range calculation algorithm

A crucial part of accurate dose delivery in proton therapy is the calculation of the range the particles travel within the irradiated tissue. While the beam range in a homogeneous medium of known compositions can be measured accurately, uncertainties are involved when patients are irradiated, which are taken into account in form of a treatment margin. While the exact size of the treatment margin varies slightly between different centers, range margins are around the value of 3.5% $\pm$ 1 mm as implemented in the Francis H. Burr proton center at Massachusetts General Hospital (MGH).

In a review paper about range uncertainties, Paganetti [9] groups the range uncertainties into two categories whether they are dependent or independent of dose calculation. The main sources of uncertainties that are dependent on dose calculation are: Biology (+0.8%), CT imaging and conversion to tissue

(0.78%), mean excitation energies (1.5%) and range degradation (up to 2.5% without MC, 0.1% with MC).

The uncertainties coming from lateral and local tissue inhomogeneities depend on the algorithm used to calculate the beam range. Lateral inhomogeneities in the irradiated material cause unaccounted multiple Coulomb scattering and range straggling, which cannot properly be modeled using simplified dose calculation algorithms as applied in most clinical planning algorithms [38, 40, 74]. The magnitude of arising uncertainties was quoted as up to 2.5%, which is the largest source of range uncertainties encountered in treatment planning. This uncertainty can be reduced by properly modeling the physics of particle interactions in tissue by using MC simulations instead of parameter-based algorithms for dose calculation. According to Paganetti, the remaining uncertainty due to inhomogeneities is 0.1% when MC is used for dose calculation. Schuemann *et al.* [39] investigated the influence of multiple Coulomb scattering on the proton range by comparing a large data set of treatment fields for different sites. They used treatment plans from a commercial treatment planning algorithm (XiO, Computerized Medical System Inc.) and recalculated them with MC to investigate the effect range degradation has on plan delivery. They found substantial differences between planning system and MC calculated beam ranges, with root mean square deviations up to 6% caused by range degradation in head and neck cases.

This chapter presents an initial study to investigate the effect of lateral inhomogeneities and the applied range calculation algorithm on the accuracy of the predicted range. We use measured WER values of animal tissue samples to investigate range uncertainties from lateral inhomogeneities by comparing predicted WER values from a) a simple ray tracing algorithm and b) MC.

3.2 Materials and methods

3.2.1 Sample preparation

In this study, 12 types of tissue samples are used. The samples are collected fresh from the butcher's on the day of the experiment and fitted into plastic containers (5 cm x 5 cm x 15 cm). The remaining space in the container is filled up with saline water, with a concentration of 9 g NaCl per liter of water. The amount of saline water that is added per container varies between 50 mL and 100 mL, depending on the sample size. Fig. 3.1 shows a picture of the samples in the container, and a list of all samples can be found in table 3.1.

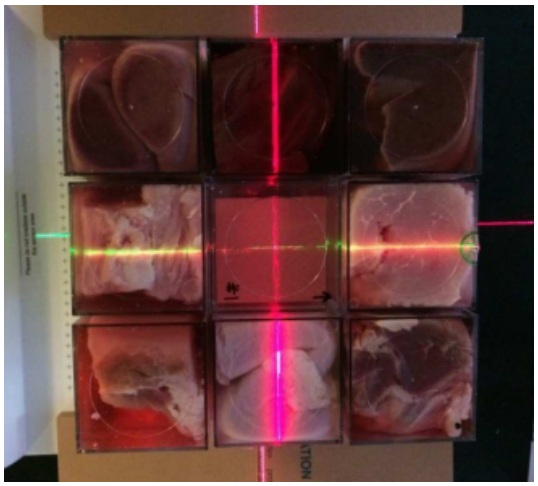


Figure 3.1: Animal organs individually packed into plastic containers, view from the top along the beam direction. After packing the organs into containers with saline water, the samples are placed in the beam line for measurements of the WER.

We measure the WER of the animal tissues using the dose extinction method, as described in the paper by Zhang *et al.* [73], and briefly illustrated here. The WER measurements are performed in a double-scattering beam line with gantry (IBA International, Louvain-La-Neuve, Belgium), at a gantry angle of 0°. A Matrixx Evolution ion chamber array (IBA Dosimetry, Bartlett, TN, USA) with 1020 ionization chambers (chamber volume 0.08 cm^3, diameter 4.5 mm, 7.62 mm center to center distance, 24.4 × 24.4 cm^2 active area) is positioned on the treatment table. For measurement of the WER, we operate the Matrixx detector in movie mode, collecting 1 frame per second, in this way

Table 3.1: List of animal tissues investigated in this study with the statistics of their ED and EAN values extracted using Bourque et al., as well as their low- and high-kVp CT numbers. The means and standard deviations are calculated over all voxels of an ROI of 308 mm^3. The ROI is placed into a single slice within the samples and is chosen to represent ED, EAN and HU of the respective tissues.

		ED		EAN		HU_l		HU_h	
	Tissue type	Mean	Std	Mean	Std	Mean	Std	Mean	Std
1	Pig stomach	1.039	0.018	7.33	0.61	43.7	10.0	42.6	11.4
2	Pig blood	1.043	0.016	7.33	0.55	48.0	7.9	47.0	9.3
3	Pig muscle	1.056	0.019	7.41	0.61	64.0	14.1	61.4	13.1
4	Cow muscle	1.057	0.020	7.30	0.61	60.8	14.1	60.6	15.1
5	Veal brain	1.036	0.019	7.37	0.69	42.5	11.8	41.2	11.6
6	Pig kidney	1.027	0.028	7.16	0.70	26.2	32.2	29.5	26.3
7	Pig liver	1.053	0.017	7.43	0.62	62.0	9.2	59.0	10.1
8	Pig leg	1.452	0.409	10.82	2.95	853.5	758.2	616.2	532.7
9	Cow tail	1.185	0.167	9.45	2.22	341.5	344.5	257.8	243.3
10	Pig rib	1.227	0.168	10.94	1.70	486.1	335.2	345.4	240.6
11	Pig vertebra	1.296	0.079	11.81	0.57	614.2	140.1	443.2	103.8
12	Pig scapula	1.262	0.314	10.14	2.68	526.6	591.7	375.8	425.0

collecting a dose profile as a function of time. The plastic containers with the biological samples are positioned on top of the detector array. A water-filled tank is placed on top of the samples. The setup is illustrated in fig. 3.2.

3.2.2 Measurement of the WER using the dose extinction method

We continuously irradiate the setup with a broad proton beam (circular field with a diameter of 25 cm). The energy of the beam is 195 MeV, corresponding to a beam range of 25 cm in water, with a modulation of 20 cm. During irradiation, we slowly drain the water from the tank, thus increasing the range of the beam in the samples. The tank and the pump are calibrated to reduce the water level by 0.4 mm per second.

The dose measured as a function of time is converted into dose as a function of water height in the column. For every detector pixel of interest, we find the water height corresponding to a dose fall-off of 80% of the maximum

dose. From this water height H_{80}, we calculate the WER of the samples as

$$\mathrm{WER}_{\mathrm{exp}} = R_{80} - H_{80} - R_{\mathrm{offset}} \tag{3.1}$$

with R_{80} the beam range as measured in the dose extinction setup, at the 80% dose fall off behind the Bragg peak. $\mathrm{R}_{\mathrm{offset}}$ includes the water tank bottom and the Matrixx build-up material, and is measured independently for this setup using a pinpoint chamber (PinPoint TN31006, PTW Freiburg, Germany). The value of $\mathrm{R}_{\mathrm{offset}}$ is 13.85 mm, and the uncertainty is determined in repeated measurements as 0.35% (68th percentile), leading to a negligible impact on WER measurements (i.e., less than 0.05 mm). The quoted uncertainty is a type A uncertainty. Type B uncertainties are not considered here but are expected to be negligible. Additionally, we account for non-uniformity of the beam across the field by applying a map of correction factors, as described in Zhang *et al.*. We find that the beam range heterogeneity is around 1.5 mm across the field. This only includes field heterogeneities, no inhomogeneities of the samples are considered at this point. The uncertainty of the resulting correction factor is estimated to be 0.2% with repeated measurements, leading to a negligible impact on WER measurements (i.e., less than 0.01 mm). Accounting for the chamber response reproducibility (±0.20%), the overall uncertainty of WER measurements with the dose extinction method is estimated to be 0.20% (68th percentile).

3.2.3 Estimation of the WER using CT

3.2.3.1 CT acquisition

To perform SECT and DECT scans of the tissues, a Siemens Somatom Flash (Siemens Sector Healthcare, Forchheim, Germany) dual-source CT scanner is used. The scan parameters for the DECT and SECT scans are summarized in table 3.2. To estimate SPRs from CT numbers, a calibration procedure must be performed. Therefore, a tissue characterization phantom Gammex RMI-467 (Gammex, Inc., Middleton, WI, USA) is scanned using the quoted

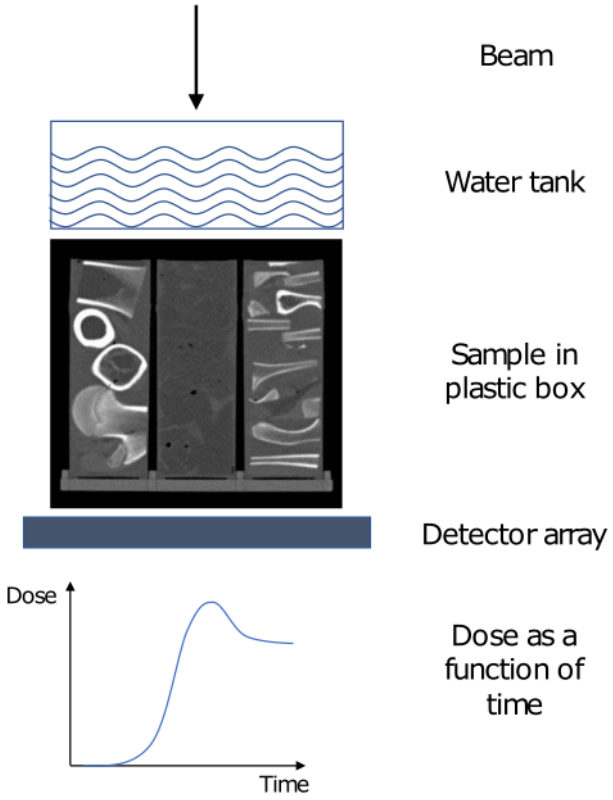


Figure 3.2: Illustration of the measurement setup for the dose extinction method. The range of the broad beam is varied by continuously decreasing the water level in the tank. The beam passes the samples and the dose behind the samples is recorded by the Matrixx detector, operated in movie mode. This way we measure the dose as a function of time. A CT slice of the samples is shown here (Window: 1500 HU; Level: 300 HU).

SECT and DECT protocols. The phantom disk has a diameter of 33 cm, the insert diameter is 2.8 cm and the insert height is 7 cm. To perform the calibration, the CT numbers of the tissue-equivalent inserts are measured in circular ROIs (226 mm^3) and averaged over 20 slices within the phantom.

Both protocols (SECT and DECT) are chosen to represent clinically applicable protocols. As a result, the DECT protocol has a higher $CTDI_{Vol}$, indicating a higher dose to the patient. We decide not to adapt SECT scan parameters to match the $CTDI_{Vol}$ of the DECT scan. Although the image quality of the SECT in terms of noise would improve, we have shown in an earlier study [70] that the SPR prediction based on Schneider *et al.* is fairly robust to noise

and an improvement in image quality would only have a minimal effect on the predicted SPR.

Table 3.2: List of scan parameters for the SECT and DECT scans. The image noise (1σ) was measured in lucite.

Parameter	SECT	DECT low	DECT high
Tube voltage	140 kV	100 kV	140 kV Sn
Exposure	300 mAs	300 mAs	232 mAs
Slice thickness	1 mm	1 mm	1 mm
Pixel width	0.46 mm	0.46 mm	0.46 mm
$CTDI_{Vol}$	29.51 mGy	23.29 mGy	23.29 mGy
Image noise (ΔHU)	9.1 HU	13.4 HU	13.0 HU
Reconstruction Kernel	B30f	B30f	B30f
CTDI phantom size	32 cm	32 cm	32 cm

3.2.3.2 CT-number-to-SPR conversion

The SECT calibration is performed using the stoichiometric method proposed by Schneider *et al.* [10]. The CT numbers of the tissue-equivalent inserts are used to find the energy fit parameters (K^{ph}, K^{coh}, K^{incoh}). With the found parameters it is possible to calculate the CT numbers of a set of human reference tissues [49]. Theoretical SPR values of these tissues are calculated from composition data using the Bethe-Bloch formula for an energy of 195 MeV. I-values of the tissues are calculated from elemental mass fractions of the tissues using the Bragg additivity rule. The elemental I-values are taken from ICRU Report 37 [27]. The calibration curve is then obtained by performing piecewise linear fits for three tissue regions (Lung: -1000 HU to -40 HU, soft tissues: -40 HU to 150 HU, bony tissues: 150 HU to 2000 HU). The calibration curve is applied voxel wise to obtain a map of SPR values.

The DECT data is processed using two different calibration methods. In a previous study [70], we evaluated the theoretical accuracy of both methods. The method by Bourque *et al.* estimates the EAN from DECT images using a parametric fit to the DEI of tissue substitutes. From the EAN estimation,

we can solve the system for the ED and estimate the I-value; both quantities are necessary to estimate the SPR using the Bethe-Bloch equation.

In the same study [70], we tested the ETD method by Lalonde and Bouchard [26], which predicts elemental compositions and electron densities of the tissues. The SPR values are calculated from elemental compositions using the Bragg additivity rule as well as the Bethe-Bloch formula. In this study, we investigate an adaptation of this method especially made for SPR prediction from noisy MECT data [47].

To calibrate the method by Bourque *et al.*, we use the CT numbers measured in the tissue-equivalent inserts to find the fit parameters b_m, as described in [13], eq. 28. From the DEI of the tissue substitutes, we find the c_k values by fitting the DEI to theoretically calculated values for the EAN, as described in [13], eq. 35. Bourque *et al.* suggest a polynomial order of 5 for the DEI-EAN fit. When single pixels are considered, it is possible that the observed DEI is outside the calibration domain due to noise. To be more robust to noise, we add two virtual materials with $Z = 3$ and $Z = 20$ to the DEI-EAN calibration. We find the DEI values of these two virtual materials by linear extrapolation from the original DEI-EAN fit. This strategy ensures the behavior of the calibration curve outside the calibration domain. The formalism by Bourque et al was investigated in the previous chapter and was found to predict RSP values of soft tissues and bones with a very low bias (0.02% and -0.77% for soft tissues and bones, respectively).

For the method of Lalonde *et al.* [47], we use the measured CT numbers of the phantom to calibrate the Z-space coefficients, as proposed in [26], eq. 24. Once the Z-space coefficients are calibrated, it is possible to decompose CT data with optimal materials following a method called Bayesian ETD. Theses optimal materials, called eigentissues, are principal components of human tis-

sue elemental compositions found in literature. By using a prior function that penalizes solutions with unlikely eigentissue fractions, Bayesian ETD extracts the maximum likelihood *a posteriori* composition in each voxel. The advantage of ETD over most DECT formalisms is the direct voxelwise estimation of elemental compositions of tissues, without performing an ED-EAN fit first or applying tissue segmentation. A suggested future application is the use of these elemental compositions for accurate MC treatment planning. However, this study aims at the validation of SPR inputs for current treatment planning systems rather than the validation of MC input parameter, which can be the subject of future studies. The formalism by Lalonde et al was investigated in the previous chapter and was found to predict RSP values of soft tissues and bones with a very low bias (0.13% and -0.08% for soft tissues and bones, respectively).

The here applied CT-number-to-SPR conversion methods are designed to predict the SPR values in human tissues. The containers, however, are made out of plastic. We ensure that the use of these methods on plastic HUs does not introduce a bias to the predicted WER. Therefore, we measure the WER of the container top and bottom in a water tank using a pinpoint chamber. The resulting WER (3.5 mm) is compared to each CT predicted WER of the container walls, and the difference is added to or subtracted from the CT predicted WER. We follow the same procedure with the trays holding the containers in place. We obtain three different maps for the WER: based on SECT images (WER_{SECT}) and based on DECT images ($WER_{Bourque}$ and $WER_{Lalonde}$). In figure 3.3, examples of the resulting 2D maps are shown for one DECT method and the measured WER map from dose extinction.

3.2.3.3 MC simulation

To take into account the inevitable beam degradation due to lateral inhomogeneities on WER calculations, MC simulations are performed using TOPAS (TOol forPArticle Simulation) version 3.1. Because MC calculations cannot be

performed directly on the SPR maps, results obtained by each method are converted to MC inputs in a two steps procedure. First, the built-in CT number to MC inputs tool of TOPAS based on the method from Schneider *et al.* [11] is used to construct a reference CT number to SPR look-up table. Then, the inverse of this look-up table is used to convert SPR volumes predicted by each method into synthetic input CT geometries. This way, we ensure that the SPR distributions seen by TOPAS correspond to the ones originally predicted by each method. The WER associated to each SPR volume is calculated using a mono-energetic proton beam of 195 MeV, with a total number of 1.2×10^7 histories. The mean energy of protons exiting the animal samples is scored in a surface detector placed at the effective point of measurement of the Matrixx detector. Conversion from proton energy to WER is done using the PSTAR database [75]. The co-registration between the MC detector and the Matrixx is performed using metallic markers routinely used in image-based treatment planning simulation, which are attached to the sample containers before the CT scan. Before each WER measurement, a radiograph of the samples and the Matrixx detector positioned on the treatment table is taken. On this radiograph, it is possible to identify the markers and thus calculate their position relative to the center of the Matrixx. This information is used to place the surface detector in the MC simulation. Since the WER is determined via the mean energy behind the samples using a reproduction of the experimental setup, it is not necessary to additionally account for the larger projected thickness of the off-axis samples.

3.2.3.4 Ray tracing

To assess the errors made by a simpler dose calculation algorithm, we estimate the WER from DECT images using a ray tracing algorithm. For this study, the SPR maps calculated from the ETD method are selected. We perform a co-registration with a virtual detector of similar proportions as the Matrixx detector. The co-registration is performed using metallic markers routinely used in image-based treatment planning simulation, which are attached to the

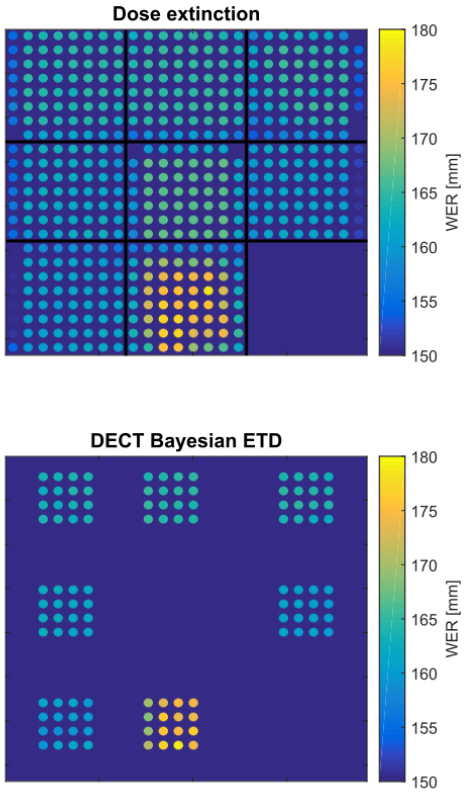


Figure 3.3: Examples of the 2D WER maps that are compared in this study. On the top are the measured WER maps using the dose extinction method; on the bottom are the WER maps obtained with MC from the ETD method. The image on the bottom indicates which chambers are used for analysis. The containers in the displayed images are filled with (from top to bottom, left to right): Pig liver, pig muscle, cow muscle, pig blood, a Lucite block, pig stomach, pig kidney, pig vertebra, pig leg bone. The pig leg bone measurement of this particular sample is not included in the analysis since it did not range out during the measurement. The measurement was repeated with less bone in the beam path.

sample containers before the CT scan. Before each WER measurement, a radiograph of the samples and the Matrixx detector positioned on the treatment table is taken. On this radiograph, it is possible to identify the markers and thus calculate their position relative to the center of the Matrixx. This information is used to place the virtual detector. The experimental setup is then simulated by performing ray tracing in plastimatch [76] through the SPR images. The grid size for ray tracing is 512×512, resulting in a pixel size of $0.48 \times 0.48\,\mathrm{mm}^2$. From ray tracing, we record the intersection length of each ray with the image voxels, as well as the corresponding SPR in this voxel. The WER along each ray is calculated by a summation of the intersection length of the ray in the voxel multiplied by the corresponding RSP. The WER values of those rays that fall on the same virtual detector element (in total 80 rays per detector element) are averaged.

3.2.3.5 Statistical analysis

For the purpose of clarifying our WER analysis, we define four statistical quantities and specify the formulas used to calculate them. For WER values obtained experimentally and predicted with a specific CT-based method, we define the WER error as follows:

$$\Delta\mathrm{WER} \equiv \frac{\mathrm{WER}_{\mathrm{CT}} - \mathrm{WER}_{\mathrm{exp}}}{\mathrm{WER}_{\mathrm{exp}}}, \tag{3.2}$$

where $\mathrm{WER}_{\mathrm{CT}}$ is the CT-based WER value (e.g., $\mathrm{WER}_{\mathrm{SECT}}$, $\mathrm{WER}_{\mathrm{Bourque}}$ or $\mathrm{WER}_{\mathrm{Lalonde}}$) and $\mathrm{WER}_{\mathrm{exp}}$ is the WER measurement with the dose extinction method as described in equation 3.1. ΔWER is determined for the number N of detector elements. From the statistical sample of N ΔWER values, we calculate the following quantities:

1. The *mean WER error*, defined as

$$\overline{\Delta\mathrm{WER}} \equiv \frac{1}{N}\sum_{i=1}^{N}\Delta\mathrm{WER}_i. \tag{3.3}$$

For each CT-based method, this quantity corresponds to the bias in WER prediction. We can show that the expectation value of the WER error is $\mathrm{E}(\overline{\Delta\mathrm{WER}}) = \mu$ (the true mean).

2. The *standard deviation of the WER error*, defined as

$$s_{\Delta\mathrm{WER}} \equiv \sqrt{\frac{1}{N-1}\left[\sum_{i=1}^{N}\Delta\mathrm{WER}_i^2 - \overline{\Delta\mathrm{WER}}^2\right]}. \tag{3.4}$$

For each CT-based method, this quantity represents the distribution spread of WER errors with respect to its mean. We can show that the expectation value of the estimator is $\mathrm{E}(s^2_{\Delta\mathrm{WER}}) = \sigma^2$ (the true variance).

3. The *root mean square error* (RMS error), defined as

$$\mathrm{RMS}_{\Delta\mathrm{WER}} \equiv \sqrt{\frac{1}{N}\sum_{i=1}^{N}\Delta\mathrm{WER}_i^2}. \tag{3.5}$$

We can show that the expectation value of the RMS error squared is $\mathrm{E}(\mathrm{RMS}^2_{\Delta\mathrm{WER}}) = \sigma^2 + \mu^2$. The square root of this quantity corresponds approximately to the limit of the symmetric interval (i.e., $[-\sqrt{\sigma^2+\mu^2}, \sqrt{\sigma^2+\mu^2}]$) containing 68% of a Gaussian distribution with mean μ and variance σ^2. Therefore, this quantity is interpreted as the overall uncertainty, i.e. $u \equiv \sqrt{\sigma^2+\mu^2}$, with a statistical significance of 68% (k=1).

We can show quite trivially that the variance of the mean error is the following quadratic sum

$$\sigma^2 = \sigma^2_{\mathrm{WER}} + \sigma^2_{\mathrm{exp}} \tag{3.6}$$

where σ_{WER} and σ_{exp} are the standard deviations (i.e., type A uncertainties) of the WER estimated with the overall CT-based method and the dose extinction method, respectively. We can also show that for each CT-based method, the

expectation value of the RMS error squared is the quadratic sum of the type A uncertainty in predicting WER with the CT method (σ_{WER}), the bias of the result (μ) and the type A uncertainty in predicting WER experimentally with the dose extinction method (σ_{exp}). That is,

$$\mathrm{E}(\mathrm{RMS}^2_{\Delta\mathrm{WER}}) = \sigma^2_{\mathrm{WER}} + \mu^2 + \sigma^2_{\mathrm{exp}}. \tag{3.7}$$

In this relation, the value of σ_{exp} is set to 0.20% and the assumption is made that the dose extinction method is unbiased. This assumption is based on results found during the validation of the workflow using tissue substitutes, see figure 3.5. In equation 3.3 we define the mean WER error as a relative quantity, we can show that the expectation value of the average error corresponds to the ratio of the expectation of the range difference over the expected range. This approximation is valid since the probability distribution of the range difference can be assumed narrow with respect to the average range and also far from the singularity of the denominator being zero. This leads to our following definition of the uncertainty of the WER estimation (with the statistical significance of k=1):

$$u_{\mathrm{WER}} = \sqrt{\mathrm{RMS}^2_{\Delta\mathrm{WER}} - \sigma^2_{\mathrm{exp}}} \tag{3.8}$$

which is an unbiased estimator of the quadratic sum of the WER standard deviation from CT only and the mean WER error, that is the expectation value of the estimator is

$$\mathrm{E}(u^2_{\mathrm{WER}}) = \sigma^2_{\mathrm{WER}} + \mu^2. \tag{3.9}$$

The uncertainty of WER estimation from CT only, u_{WER}, includes the uncertainties from range degradation due to multiple Coulomb scattering u_{deg}. The term range degradation refers to a widening of the distal dose fall-off of the Bragg peak, originating from increased multiple Coulomb scattering at tissue interfaces and inhomogeneities. Also included into u_{WER} are the uncertainties

from the I-value u_I and the uncertainty from the CT calibration an conversion u_{CT}. Hence we can define

$$u^2_{\mathrm{WER}} = u^2_{\mathrm{deg}} + u^2_I + u^2_{\mathrm{CT}}. \tag{3.10}$$

The contribution from range degradation u_{deg} for soft tissues is assumed to be negligible since no soft tissue-bone interfaces are present in the samples. For the bones, we assume an uncertainty of 0.14%, based on the value reported by Paganetti (2012).

The contribution of the I-value uncertainty u_I to range uncertainty was estimated previously [28, 32, 77, 9]. In this study, we expect to observe a combined uncertainty of the contributions from I-value and CT calibration and conversion, defined as

$$u^2_{\mathrm{comb}} = u^2_I + u^2_{\mathrm{CT}}. \tag{3.11}$$

3.2.4 Validation of the workflow using Gammex RMI-467 tissue substitutes

Our implementation of the dose extinction method is validated using tissue substitutes from a Gammex RMI-467 phantom. The reference SPR values are measured using a pinpoint ionization chamber (PinPoint TN31006, PTW Freiburg, Germany) in a broad beam (25 cm range, 20 cm modulation). We use 12 tissue-equivalent inserts of the Gammex RMI-467 phantom, as listed in table 3.3. To measure the WER of the Gammex inserts with the dose extinction method, we remove the inserts from the phantom disk and place them on the Matrixx. In this work, we additionally validate our SECT and DECT calibrations using the SPR values from the pinpoint chamber. After calibrating SECT and DECT with the Gammex RMI-467 phantom, we apply the calibration to the phantom images and measure the SPR throughout the phantom inserts. The SECT calibration curves for tissues and substitutes as

well as the curves to convert the EAN into I-values are displayed in figure 3.4. It should be noted that the SECT calibration curve for this validation study is not established using the method by Schneider *et al.* (1996). To establish a calibration curve, we apply a piecewise linear fit between measured CT numbers in the Gammex RMI-467 phantom inserts and measured reference SPR. This curve does not represent human tissues but is suitable for the use in tissue-equivalent plastics. The displayed results for SECT hence represent the errors coming from the fit procedure itself. Similarly, we proceed with the DECT method by Bourque *et al.*. This method uses a polynomial fit to relate the EAN to I-values based on the idea of Yang *et al.* [23]. The differences between the approaches are the definitions of EAN and the use of continuous function instead of a piecewise fit. In their paper, Yang *et al.* suggest to perform the fit based on theoretical I-values calculated from Woodard and White [49] composition data, representing human tissues. Hence, to make this method suitable for the use in tissue substitute materials, we perform the fit based on I-values calculated from compositions of the substitutes. We do not adapt the method of Lalonde and Bouchard to plastic materials as the whole method is designed to describe human tissues.

3.2.5 Comparison of two range calculation algorithms

We compare the predicted WER (ray tracing and MC) to the measured WER values. Data points that contain a large portion of the container wall, air or mixtures of tissue and air are excluded. It follows that for each tissue container we can include 16 detector elements into our analysis. This way we obtain 112 measurements for the seven investigated soft tissues, and 80 measurements for the five investigated bone tissues.

Table 3.3: List of tissue-equivalent materials used for validation of our measurement techniques. Material-specific SPR values are given for: i) a reference measurement with a pinpoint chamber, ii) the dose extinction measurement, ii) the SECT calibration and iv) the DECT calibration by Bourque *et al.*. The uncertainty of the reference SPR values is estimated from repeated measurements to be 0.20% and the uncertainty of SPR values measured with the dose extinction method is estimated to be 0.20%. Uncertainties of SECT and DECT are estimated by calculating the standard deviation of the mean over an ROI of 243 mm^3.

	Tissue-equivalent material	SPR Reference	SPR dose extinction	SPR SECT	SPR DECT
1	Water	1.000	1.000	1.006 (0.001)	0.998 (0.002)
2	LN300 Lung	0.286 (0.001)	0.286 (0.001)	0.287 (0.001)	0.275 (0.006)
3	AP6 Adipose	0.951 (0.002)	0.950 (0.002)	0.941 (0.001)	0.943 (0.002)
4	BR12 Breast	0.978 (0.002)	0.977 (0.002)	0.958 (0.001)	0.971 (0.002)
5	Solid Water	1.004 (0.002)	1.005 (0.002)	1.007 (0.001)	0.996 (0.002)
6	LV1 Liver	1.078 (0.002)	1.080 (0.002)	1.088 (0.001)	1.080 (0.002)
7	SR2 Brain	1.069 (0.002)	1.068 (0.002)	1.050 (0.001)	1.063 (0.002)
8	CB2 - 30% $CaCO_3$	1.267 (0.003)	1.266 (0.003)	1.232 (0.001)	1.262 (0.003)
9	CB2 - 50% $CaCO_3$	1.431 (0.003)	1.428 (0.003)	1.421 (0.001)	1.435 (0.002)
10	SB3 Cortical Bone	1.621 (0.003)	1.614 (0.004)	1.636 (0.001)	1.628 (0.002)
11	B200 Mineral Bone	1.095 (0.002)	1.100 (0.002)	1.106 (0.001)	1.099 (0.002)
12	IB3 Inner Bone	1.076 (0.002)	1.076 (0.002)	1.096 (0.001)	1.073 (0.003)

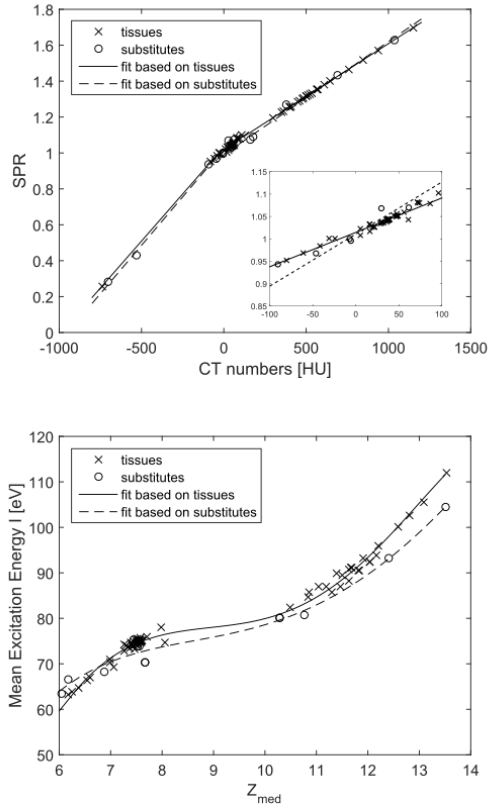


Figure 3.4: Calibration technique on tissue-equivalent plastics: On the top, calibration curves for the SECT methods are shown, where the solid line represents the fit for the human reference tissues and the dashed line represents the fit for tissue substitutes; the soft tissue region is shown as inlet; on the bottom, the curves to convert EAN into I-values for the Bourque method, where the solid line represents the fit for the human reference tissues and the dashed line represents the fit for tissue substitutes.

3.3 Results

3.3.1 Validation of the workflow using Gammex RMI-467 tissue substitutes

In the validation measurement, we show good agreement between SPRs measured by dose extinction and our reference SPR with an RMS error of 0.29%. We consider this value as part of the overall uncertainty of the experimental method to determine WER. The resulting differences in SPR values determined with dose extinction, SECT and DECT are summarized in fig. 3.5. Table 3.4 summarizes the mean error and standard deviation as well as RMS

errors on SPR for each investigated method. Our results clearly demonstrate the limitations of the use of a standard SECT calibration curve in agreement with earlier studies [15, 72]. We report an RMS error on SPR of 1.59% between SECT-determined values and reference SPR. For DECT, we report RMS errors of 0.61%. This value is in accordance with the original values published in Bourque *et al.*, who report an RMS error of 0.67% (for the same energy couple) on experimentally determined SPR values.

Table 3.4: Mean error, standard deviation and RMS error in % between reference SPR values (pinpoint chamber), dose extinction measured SPR values and SECT and DECT predicted SPR values. Quoted values are relative to water.

Method	Mean	Std	RMS error
Dose extinction	-0.06	0.28	0.29
SECT	-0.25	1.57	1.59
DECT $\rho_e - Z$	-0.28	0.55	0.61

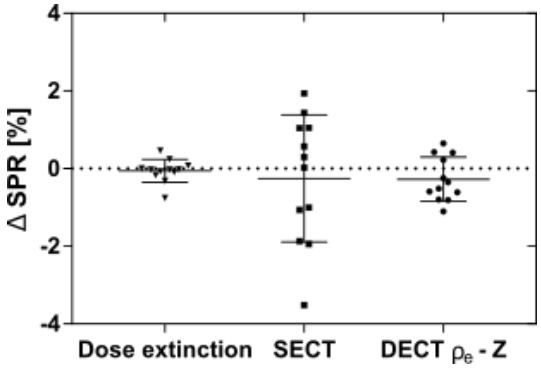


Figure 3.5: Validation of our measurement methods. We compare SPR values from tissue-equivalent materials of the Gammex RMI-467 phantom determined with 1) dose extinction (left), 2) an empirical SECT calibration curve (middle) and 3) the stoichiometric DECT calibration by Bourque *et al.* (right) to the SPR measured in a water tank. Quoted values are relative to water. Also presented are mean errors and the standard deviations.

3.3.2 Tissue samples

For each investigated tissue type, we collect 16 data points from adjacent ionization chambers of the array, presented in figures 3.6 and 3.7. With 16 data points per sample, we extract mean error and standard deviation for the

WER per organ, resulting in a total number of detector elements of $N = 112$ for soft tissues and $N = 80$ for bones. To determine if the WER data come from a normally distributed population, we apply a Lilliefors test. The test result showed that all data with few exceptions are normally distributed, therefore we can apply a paired two-tailed t-test on most of the data. For the not normally distributed exceptions, we apply a Wilcoxon signed-rank test. We perform the appropriate statistical test on each pair of observed WER differences (SECT vs. DECT $\rho_e - Z$, SECT vs. DECT ETD, DECT $\rho_e - Z$ vs. DECT ETD). For most pairings, we observe significant differences ($p<0.05$) between WER predictions. Whenever two distributions are not found to be significantly different, it is explicitly indicated in the graphs of figures 3.6 and 3.7. Additionally, table 3.5 shows an overall uncertainty on the WER estimation. We observe uncertainties of 0.53% and 1.37% for SECT (soft tissues and bones respectively), 0.19% and 1.06% for DECT using the method by Bourque *et al.* (DECT $\rho_e - Z$) and 0.38% and 1.06% for the DECT ETD method. A paired t-test is performed on the overall distributions of WER errors for soft tissues and bones, showing significant differences ($p<0.05$) for all pairings.

Table 3.5: Statistics of the WER errors in % for soft tissues and bones from CT imaging only. The bias is the mean WER error (Eq. 3.3). The type A uncertainty is the standard deviation of the WER error (Eq. 3.4) with a quadratic subtraction of the experimental uncertainty (σ_{exp}=0.20%). The overall uncertainty includes range degradation, I-value and CT calibration, and it is equal to the quadratic sum of the mean WER error and the standard deviation of the WER error (Eq. 3.8). Uncertainties are reported with a statistical significance of k=1.

	Soft tissues			Bones		
Method	Bias	Type A	Overall	Bias	Type A	Overall
SECT Schneider *et al.*	-0.44	0.29	0.53	-0.88	1.06	1.37
DECT $\rho_e - Z$	-0.01	0.19	0.19	-0.58	0.89	1.06
DECT ETD	-0.33	0.20	0.38	-0.14	1.05	1.06

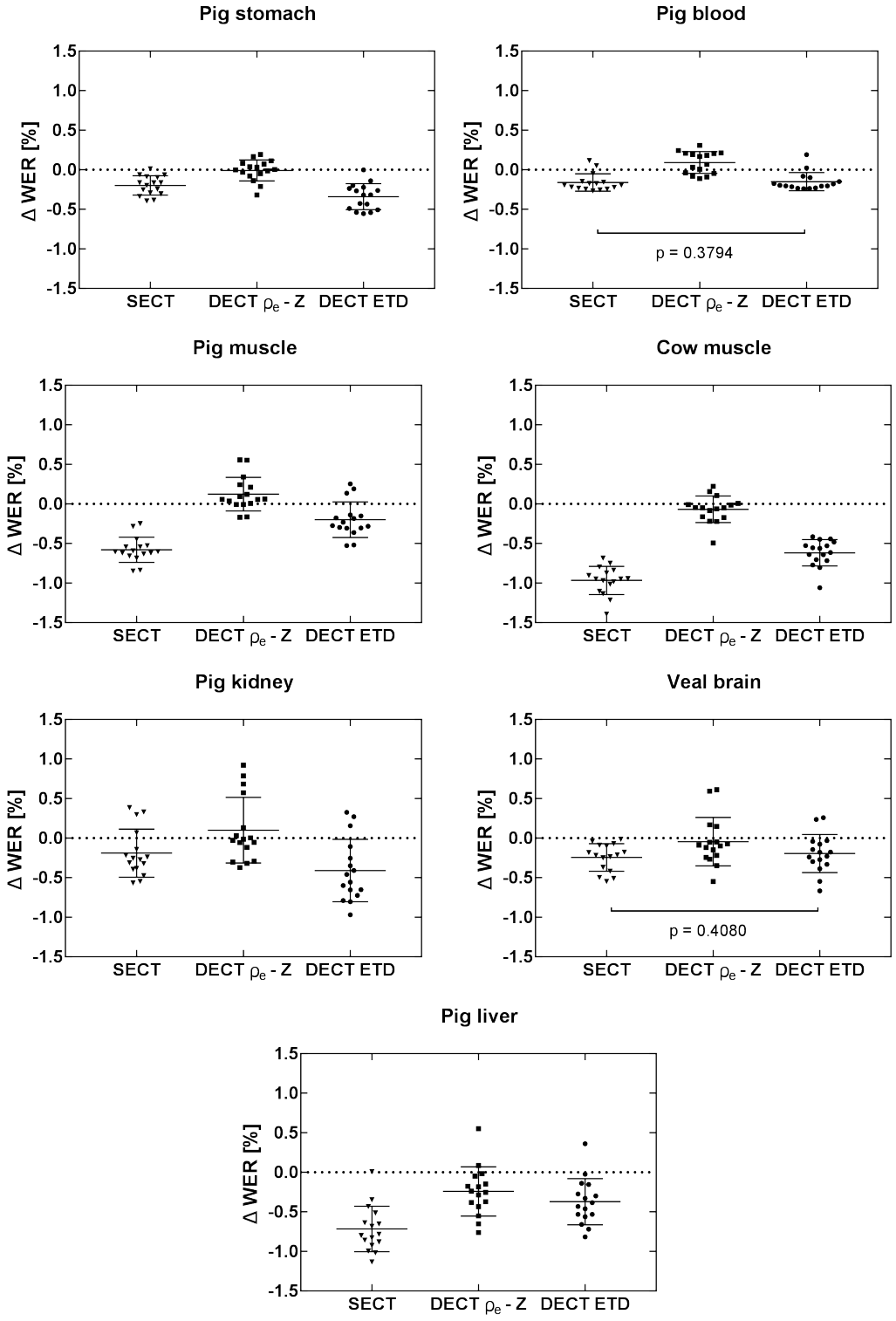


Figure 3.6: Soft tissue samples: The tissue-specific percentage difference between the WER determined with dose extinction and CT imaging is shown here. For every tissue type, the WER differences, mean error and standard deviation are shown for the stoichiometric SECT calibration (Schneider *et al.*), the stoichiometric DECT calibration (Bourque *et al.*) and the DECT ETD method (Lalonde *et al.*). The observed WER distributions are all significantly different with two exceptions (SECT vs. DECT ETD for pig blood and veal brain).

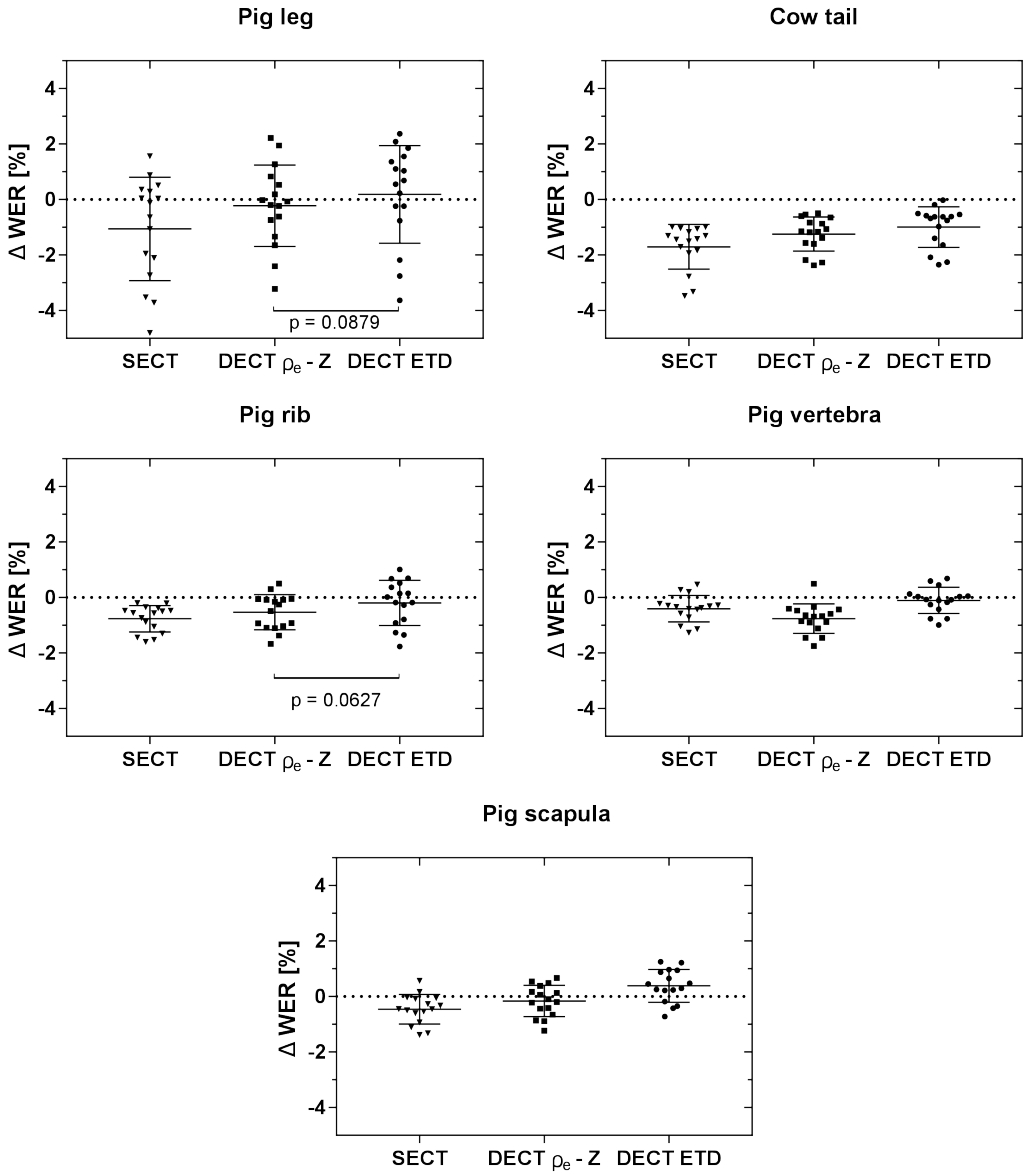


Figure 3.7: Bone samples: The tissue-specific percentage difference between the WER determined with dose extinction and CT imaging is shown here. For every tissue type, the WER differences, mean error and standard deviation are shown for the stoichiometric SECT calibration (Schneider *et al.*), the stoichiometric DECT calibration (Bourque *et al.*) and the DECT ETD method (Lalonde *et al.*).

3.3.3 Estimation of uncertainties in CT-number-to-SPR conversion

A detailed uncertainty budget is performed based on the statistical behavior of the data and estimations of experimental uncertainties and uncertainties from range degradation due to lateral inhomogeneities. In tables 3.6 and 3.7, we summarize the uncertainties involved in the WER estimation arising from CT calibration and I-value. We establish an uncertainty budget for soft tissues and bones separately, due to differences in u_{deg} for the two different tissue types. The uncertainties arising from range degradation as well as I-value occur in both imaging modalities (i.e., SECT and DECT). The differences in the observed uncertainties u_{WER} for SECT and DECT must, therefore, arise from the CT calibration and conversion technique and I-value. In soft tissues, we find an overall difference of -0.34% (68th percentile) between u_{comb} of DECT and SECT. In bones, we find an overall difference of -0.31% (68th percentile) on u_{comb} for DECT over SECT, which is mostly dominated by the bias observed with the SECT method. The results are summarized in Tab. 3.8.

The voxelwise prediction of the SPR using SECT and DECT is subject to CT imaging artifacts and errors from the CT-number-to-SPR conversion. These errors and resulting range uncertainties were investigated within a theoretical scope in our previous work [70]. For noise levels comparable to this study (7 HU for SECT, 8 HU and 12 HU for DECT), we reported a theoretical error on the range (95th percentile) in soft tissues of 0.68% for SECT and 0.30% for DECT. These values are taken from figure 4a) in [70]. Based on the methods applied in the previous study, we estimate the theoretical range uncertainties to compare to the findings of the present study. The theoretical range uncertainty estimation for soft tissues is done for 16 cm of WER to match the thickness of the samples. For bones, we quadratically combine the uncertainty from 8 cm of bones and 8 cm of soft tissues. In soft tissues, we find a difference of -0.34% (68th percentile) between u_{CT} of DECT and SECT. In bones, we find a difference of -1.14% (68th percentile) on u_{CT} for DECT over

SECT, which is mostly dominated by the bias observed in the SECT method. The results are summarized in Tab. 3.8.

Table 3.6: Uncertainties of estimated WER in % over all soft tissues. The uncertainty due to beam degradation from multiple Coulomb scattering is assumed negligible for soft tissues. The combined uncertainty on CT calibration and conversion and I-value u_{comb} is found by quadratic subtraction (Eq. 3.10).

Soft tissues	u_{WER}	u_{deg}	u_{comb}
SECT Schneider *et al.*	0.53	negligible	0.53
DECT $\rho_e - Z$	0.19	negligible	0.19
DECT ETD	0.38	negligible	0.38

Table 3.7: Uncertainties of estimated WER in % over all bones. The uncertainty due to beam degradation from multiple Coulomb scattering is estimated from Paganetti (2012) [9]. The combined uncertainty on CT calibration and conversion and I-value u_{comb} is found by quadratic subtraction (Eq. 3.10).

Bones	u_{WER}	u_{deg}	u_{comb}
SECT Schneider *et al.*	1.37	0.14	1.36
DECT $\rho_e - Z$	1.06	0.14	1.05
DECT ETD	1.06	0.14	1.05

Table 3.8: Comparison of the uncertainty from imaging (u_{comb}) of the method by Bourque *et al.* with theoretical data. Theoretical data are estimated using the method of Bär *et al.*[70]. Theoretical data are not affected by the uncertainty on I-values ($u_I = 0$).

	Experimental			Theoretical		
	SECT	DECT	Difference	SECT	DECT	Difference
Soft tissue	0.53	0.19	-0.34	0.55	0.21	-0.34
Bones	1.36	1.05	-0.31	1.76	0.62	-1.14

3.3.4 Comparison of two range calculation algorithms

The resulting mean WER errors per tissue type for both, MC and RT, are displayed in figure 3.8. In soft tissues, we observe a general overestimation of the measured WER using the simple ray tracing algorithm, while MC seems to slightly underestimate the measured WER. In bones, we find that ray tracing underestimates the range for most of the samples while MC reduces this bias. Using RT, the mean errors and standard deviations on WER are 0.50±0.52%

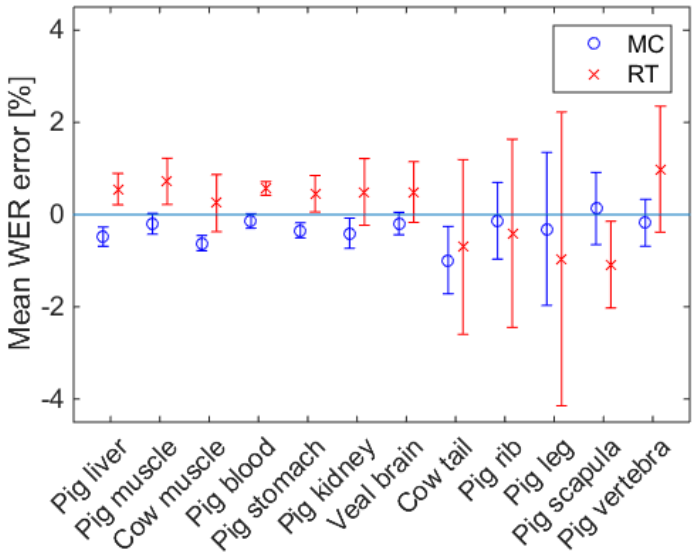


Figure 3.8: Mean WER error and standard deviation per examined tissue. For each tissue, 16 WER values were measured at different positions during the dose extinction experiment and compared to the WER calculated using ray tracing and MC.

for soft tissues and -0.43±2.12% for samples containing bones. With MC, these values are reduced to -0.34±0.27% in soft tissues and -0.30±1.03% in samples containing bones. The RMS error over all samples is 2.75% for ray tracing and 1.41% for MC, suggesting a reduction in WER calculation errors of 1.35%. For ray tracing, we observe a minimum error of -6.09% and a maximum error of 3.44%, which can be reduced to -3.63% and 2.37% with MC. The mean error, standard deviation and RMS error for each tissue are listed in table 3.9. Figure 3.9 shows a histogram of the absolute percentage WER prediction error for both methods taking all samples into account. We apply a Welch's t-test to the population to determine the statistical significance of the results. The test rejects the null hypothesis that the data come from normal distributions with equal means ($p<0.05$).

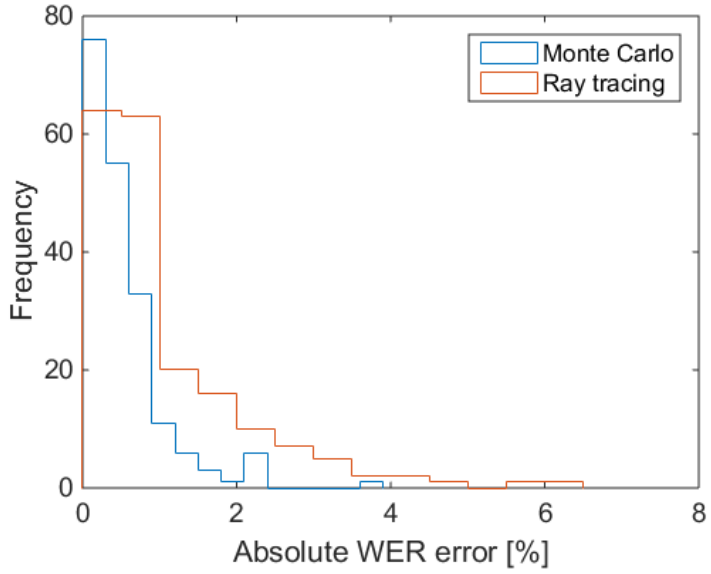


Figure 3.9: Histograms of the absolute WER error predicted by MC (blue) and ray tracing (red) over all samples and all points of measurements.

Table 3.9: Statistics of the WER prediction errors from both range calculation algorithms. For each tissue, 16 points of measurement were taken at different positions in the sample.

	Ray tracing			MC		
Tissue	Mean WER error	std	RMS error	Mean WER error	std	RMS error
Pig liver	0.55	0.34	1.06	-0.48	0.21	0.85
Pig muscle	0.72	0.50	1.43	-0.20	0.22	0.49
Cow muscle	0.25	0.62	1.07	-0.62	0.17	1.05
Pig blood	0.57	0.15	0.95	-0.14	0.16	0.33
Pig stomach	0.45	0.40	0.96	-0.34	0.16	0.61
Pig kidney	0.49	0.72	1.38	-0.40	0.33	0.83
Veal brain	0.49	0.66	1.38	-0.19	0.24	0.53
Cow tail	-0.70	1.89	3.69	-0.99	0.73	2.27
Pig rib	-0.41	2.04	3.60	-0.13	0.83	1.45
Pig leg	-0.96	3.19	6.28	-0.31	1.66	3.18
Pig scapula	-1.08	0.94	2.59	0.13	0.78	1.41
Pig vertebra	0.99	1.37	2.89	-0.18	0.51	0.92

3.4 Discussion

3.4.1 Performance of DECT methods in determining SPR

In this study, we measure the WER of biological samples containing several soft and bony tissues from three different types of animals. We pack isolated organs (when possible) in plastic boxes, filled up with saline water. We

measure the WER of a proton beam through the samples using the dose extinction method with the Matrixx detector. We scan the samples with SECT and DECT and determine the SPR voxelwise from three different calibration methods. We then predict the WER of the proton beam through the sample by performing MC simulations on the CT-predicted SPR maps. Our measurement method is validated against experimental reference SPR values of tissue-equivalent materials. The SPRs determined with the dose extinction method show excellent agreement with their respective reference SPRs. The reference SPRs determined experimentally are then used to validate the CT-based SPR values. SPRs of tissue-equivalent materials predicted from DECT show a much better agreement with reference values than SPRs predicted by SECT.

Our results with animal samples show that overall DECT performs well in determining SPRs of biological samples. The errors on WER prediction using SECT reported herein agree partially with earlier studies [50, 73]. The WER of soft tissues can be predicted well within 1% uncertainty, but higher uncertainties for bones are observed, mostly due to the bias observed with the SECT technique. Range errors in bones are lower than the results reported by Schaffner and Pedroni [50]. They predicted an overall error in range, caused by CT artifacts and CT calibration, of 1.1% in soft tissues and 1.8% in bones. It should be mentioned here that the range uncertainty for bones in their study was based on a single piece of bone scanned in air, for which they adapt a special calibration curve. Since their bone samples were scanned in air, they also scanned their calibration materials in air and adjusted the calibration curve accordingly. This procedure establishes similar beam hardening conditions for calibration and scanning and therefore reduces the bias in bone tissues. In the present study, we use a single calibration curve for all tissues to reproduce clinical conditions. Additionally, we used five different kinds of bones from two different animals, and take 16 measurements per piece of bone.

The stoichiometric DECT calibration by Bourque *et al.* improves range predictions by 0.34% in soft tissue and 0.31% in bones. Figure 3.6 shows that the SPR prediction from Bourque *et al.* reduces the bias in range prediction for each of the investigated soft tissues. For the bony tissues, we observe a reduced bias for 4 out of 5 investigated tissue types, with the exception of pig vertebra. Even though there is no improvement observed for this particular sample between SECT and DECT, the bias for both techniques is very low.
As a third CT calibration method, we apply the ETD method as proposed by Lalonde *et al.*. We observe improvement in range predictions by 0.15% for soft tissues and 0.31% for bones. From figure 3.6 we observe that for three soft tissue types, the ETD method outperforms SECT based range predictions. For two tissue types, we observe a performance similar to SECT. For two tissue types, we observe that SECT predicted ranges are slightly better than ETD predicted ranges. It is worth noting that the ETD method is outperformed by the $\rho_e - Z$ formalisms in soft tissues. One hypothesis is that the higher concentration of salt compared to biological tissues in the saline solution affected the accuracy of the ETD method, designed specifically for human tissue compositions. This behavior is expected due to the penalizing of unlikely solutions in the Bayesian approach of the ETD, which is used in this study. The prior function is constructed from human tissue composition, which makes the method potentially less applicable to tissues different from human tissues. It is however possible to adjust the severity of the regularization function. In this study, we used a regularization value of $\alpha = 0.4$, as suggested in the original publication of Lalonde *et al.*. However, it would be possible to optimize this parameter to reduce the severity of regularization, which can increase suitability of the method to tissues with compositions different from those of human tissues. It should be mentioned here that the main advantage of this method, contrarily to the two others, is that the ETD method is also suitable to estimate any other physical quantity relevant for proton therapy because

it estimates elemental compositions prior to calculate the SPR. One possible application would be to use the estimated compositions to define the materials used for MC simulations. In this study, MC input data are obtained from SPR maps using a look-up table. This is one way of defining input materials, although not necessarily optimal. Future work can focus on implementing the direct use of DECT predicted elemental compositions, as suggested in earlier studies [24, 71, 25]. Finally, one other advantage of the ETD formalism is its generality, that makes it suitable for any number of CT energy. Using more sophisticated detectors, such as photon counting detectors (PCD), the method could potentially yield a lower bias on range prediction.

3.4.2 Improvements from SECT to DECT

This study is the first to validate SECT and DECT predicted WERs in a variety of animal soft tissues and bones using MC to simulate the proton path. One major finding of this study is the validity of the SECT calibration in the investigated soft tissues. SECT and DECT both predict the experimentally measured ranges with a high accuracy level of 1% or better, and neither method is largely biased. This finding validates that a SECT stoichiometric calibration curve is well suited for the use in soft tissues. Future work could focus on the investigation of adipose tissue, which is not considered in this study.

In samples containing bones, we observe an overall improvement of 0.34% with DECT. These results indicate that one major gain in the use of DECT over SECT lies in the bony tissues. Before concluding on the validity of a DECT method and recommend it for clinical use, it's benefit in bones needs to be validated, especially since SECT performs very well in soft tissues and no major advantages are to be expected. Furthermore, intensive studies need to be performed taking lung tissues into account.

We attempt to provide a detailed uncertainty budget to isolate the uncertainty coming from tissue characterization from SECT and DECT imaging

only. Besides the uncertainties of the WER measurement, we estimate the sources of uncertainties from beam range degradation due to lateral heterogeneities, the I-value, as well as the CT imaging and calibration methods (i.e., conversion of CT data to SPR). Our soft tissue samples are homogeneous, hence one expects fewer partial volumes artifacts and negligible effects from range degradation due to Coulomb and nuclear scattering, which is more relevant in bones. Here, we use a constant uncertainty value of 0.14% to quantify the remaining uncertainties from lateral inhomogeneities. The uncertainty on SPR arising from the I-value is expected to be smaller in soft tissues than in bones due to the significant water content. Uncertainties reported in Tab. 3.6 and 3.7 are based on approximate estimations of uncertainty contributions from range degradation, however we report a combined uncertainty for CT imaging, CT calibration and I-value.

For water, an I-value of 75.3 eV is used. This value is derived using the Bragg additivity rule, elemental I-values for hydrogen and oxygen are taken from the ICRU report 37. We choose to comply with the Bragg additivity rule for consistency reasons since the DECT method by Lalonde and Bouchard derives elemental compositions and uses the Bragg additivity rule to calculate composite I-values. A different I-value for water can potentially result in a systematic shift of the CT-predicted WER values. However, the impact of a different I_{water} on the SPR and therefore the WER uncertainty depend on the tissue type, since correlations in soft tissues need to be taken into account.

The uncertainties we quote here are the combined uncertainties from CT imaging and calibration and I-values. The uncertainty on the I-value was determined in earlier studies by Bichsel and Hiraoka [32] and Kumazaki *et al.* [28], quoting a contribution on range uncertainty of 1.5% [9]. Our results however suggest that this number is overestimated, especially in soft tissues. As mentioned before, it can be shown mathematically that the contribution

of I-value uncertainties on range uncertainty depends in the tissue type. As De Smet *et al.* [78] show, the uncertainty on SPR originating from I-value uncertainties can be as low as 0.1%, depending on the tissue's water content. Another study by Yang *et al.* reports similar results, quoting SPR uncertainties of 0.23% (soft tissues) and 0.65% (bones) from uncertainties in I-values [51]. Hence, range uncertainties in soft tissues are much lower than range uncertainties in bones, which is in agreement with what we observe in this study.

In this study we quote the improvement from SECT to DECT as a difference in the combined uncertainty. We expect that the uncertainty arising from the I-value (u_I) is larger than the uncertainty arising from CT imaging and calibration (u_{CT}). Consequently, it can be shown mathematically that the uncertainty difference arising from CT imaging and calibration only must be higher than the combined uncertainty difference. Hence our quoted values represent the lowest possible improvement in (u_{CT}) that can be achieved using DECT over SECT, when I-value uncertainties are taken into account. This is in agreement with our theoretical study. Results in Tab. 3.8 show that the experimental results for soft tissues are in good agreement with theoretical estimations of range uncertainties, suggesting overall improvements on SPR estimations with DECT. For bones, we observe smaller improvements experimentally than theoretically. However, the improvements in the theoretical study are only quoting improvements from CT imaging and calibration and are hence expected to be larger than the combined uncertainty difference. This suggests that further work is necessary to achieve accurate information on the uncertainty of the I-value and evaluate the effects of using the Bragg additivity rule to determine I-value on the proton beam range errors.

Additional sources of uncertainties could also arise from our setup. Although our experimental setup is reproducible since we use organs tightly packed in plastic containers, a slight movement of tissues within the con-

tainers during the transport between CT and treatment room are difficult to avoid. Furthermore, the representation of human tissues using animal tissues has some limitations. While we observe very similar ED and EAN in soft tissues, animal bones can be denser than human bones. For this reason, we deliberately avoided the use of very dense animal bones such as a cow's leg. The samples were fresh from the butchers, collected on the day of the experiment. The preparation time was kept to a minimum to ensure that tissue characteristics as close to an in-vivo situation as possible. An additional CT scan after the WER measurement was performed for each sample to check if tissue properties had changed during the experiment. Analysis of those data showed no significant change between scans. Hence, it seems reasonable to assume that results are not strongly influenced by the sample preparation. However, the saline water may have an impact on the CT predicted WER, which will not be seen in an in-vivo situation since salt concentrations are much lower.

The range measurements were performed in a scattered beam with spread out Bragg peak, mandated by the available equipment. Here presented results address the accuracy of tissue characteristics as estimated from SECT and DECT. Those quantities are assumed to be independent of the delivery technique, hence validity of the reported uncertainties in spot scanning delivery techniques can be assumed. It is recommended to validate any DECT tissue characterization technique before it is used for treatment planning. The presented measurement techniques are easily adoptable to spot scanning techniques, e.g. by using water column measurements instead of dose extinction.

3.4.3 Comparison of range calculation algorithms

Proton range uncertainties arising from lateral inhomogeneities and the influence of the range calculation algorithm are investigated. Therefore, a set of measured WER from animal tissue samples with inhomogeneous structures is used. For comparison, the WER is predicted based on a DECT scan of the

samples using two different ways to calculate the beam range, ray tracing and MC, which are then compared to the measured WER.
For ray tracing, we observe a systematic overestimation of soft tissues in comparison to the measured WER. The same algorithm underestimates the majority of bone tissues. In contrast to that, MC shows a slight systematic underestimation of all tissues but is overall closer to the measured value than ray tracing. As expected, both range calculation methods give different results for the WER. This behavior might arise from the difference between RSP summation (as done in ray tracing) and integral from the energy loss (MC).

Generally, high errors in range prediction of up to -6.09% are observed using ray tracing, especially in inhomogeneous tissues. These results are in agreement with the results obtained by Schuemann *et al.*, who found dose differences of up to 6% between a pencil beam algorithm (XiO) and MC predicted ranges in head and neck patients. In a clinical situation, such large-scale errors in range predictions would not be covered by the clinically applied margin and might lead to an erroneous treatment. This result suggests that margins should be chosen carefully and site-specifically.

As expected, an improvement in using MC to calculate the beam range is observed. The use of MC reduces the bias in range prediction for both, soft tissues and bones, due to a more accurate modeling of the physical processes. From our experimental data, we can conclude a reduction in RMS error for the WER prediction of 1.35% over all tissues.

It is worth noting that MC reduces the standard deviation of the predicted ranges in all cases except pig blood. These results were expected since pig blood is reasonably close to water in density and composition and very homogeneous, hence ray tracing can accurately predict the WER.

The error reduction with MC as compared to ray tracing is especially pronounced in the highly heterogeneous bone samples. We find a lower mean prediction error with MC, and the standard deviation is considerably lower using MC, with 2.12% with ray tracing and 1.03% for MC.

For this study, the SPR maps calculated from the ETD method are selected. Initially, the ETD method was selected to compare the algorithms because it can provide direct input data in form of material density and composition which should improve the MC approach. In this study, however, the SPR maps are used and the exploration of advantages of using density and compositional data for MC calculations are left for future studies.

Those results show the absolute necessity for accurate range prediction and suggest the use of MC for dose calculation of proton therapy treatment. Current arguments against using MC clinically are the high computational complexity and the associated long computing times for full treatment plans. Additionally, efforts need to be made to improve input data for MC dose calculations. Available MC platforms (TOPAS as well as commercialized systems) extract their inputs from a SECT scan of the patient. The acquired CT numbers are segmented into tissues and assigned a reference elemental composition, while the density is scaled by the corresponding RSP. This procedure introduces uncertainties since it neglects tissue variability. Possible improvements with DECT and MECT are currently under investigation.

This part of the work presents an initial study aimed at understanding and quantifying how MC can improve proton range prediction, and if possible reductions in uncertainty margins can be achieved. One limitation in the study as currently performed is the difference in exit energies between the experiment and the MC simulation. It can be assumed that the severity of range degradation changes with energy, because scatter increases with decreasing

energy. In the MC simulations, the entrance energy of the proton beam is kept at a constant value. The residual energy of the beam after passing through the samples is scored and used to find the WER. In the experiment, however, we continuously increase the entrance energy until the Bragg peak was seen by the detector. Hence, the average energy as seen in the detector is always close to zero. This mismatch of energies might introduce errors in the MC simulations. Although they are not expected to be large, it might be worth investigating if this could be the origin of the observed bias.

3.5 Conclusion

This study aims at validating two DECT formalisms to extract SPR values using heterogeneous animal tissue samples. SECT and DECT scans of homogeneous tissue-equivalent plastic materials are used to estimate their SPRs and the same technique is applied on the biological samples to predict their WER in MC simulations. Using WER measurements with the dose extinction method, we isolate an uncertainty estimation of combined CT imaging artifacts, CT-number-to-SPR conversion and I-value on the range of a proton beam. We observe clear improvements in determining SPRs with DECT for homogeneous tissue substitutes as well as improvements in determining WER in animal tissues since range errors are dominated by the bias produced by the SECT method. We conclude benefits of 0.34% in soft tissues and 0.31% in WER prediction in the presence of bony tissues using DECT over SECT. Future MC-based treatment planning systems, however, might further benefit from improved input data derived with the ETD method.

Overall, chapters 2 and 3 demonstrate that DECT can provide additional information and therefore has the potential to increase accuracy in the prediction of tissue parameters for proton radiotherapy. Both, the theoretical and experimental study agree that uncertainties from CT imaging and conversion to tissue can be reduced by about 0.3% by choosing DECT over SECT.

We compare water equivalent ranges of a proton beam traveling through a set of homogeneous and heterogeneous animal tissue samples with the WER predicted by two different range calculation algorithms. We observe high prediction errors of up to -6.09% using ray tracing, which can be reduced by a factor of about two using MC. This work demonstrates the gain in range accuracy when using MC for dose calculation in the presence of tissue inhomogeneities. A reduction of mean range errors and standard deviations can be achieved in both, homogenous soft tissues and inhomogeneous soft tissue-bone mixtures.

Chapter 4

Optimized I-values for the use with the Bragg additivity rule and their impact on proton stopping power and range uncertainty

The work presented in this chapter can be found in the following journal article:

1. **Bär, E.**, Andreo, P., Lalonde, A., Royle, G., Bouchard, H. (2018) Optimized I-values for the use with the Bragg additivity rule and their impact on proton stopping power and range uncertainty. *Phys. Med. Biol., in press.* doi: 10.1088/1361-6560/aad312 [79].

The work was presented at the following meetings:

2. **Bär, E.**, Andreo, P., Lalonde, A., Royle, G., Bouchard, H., (2017) A new method to improve the accuracy of the Bragg additivity rule for calculating proton stopping powers. London, *4th PPRIG Workshop.*

3. **Bär, E.**, Andreo, P., Lalonde, A., Royle, G., Bouchard, H., (2018) A novel method to estimate mean excitation energies and their uncertain-

ties for particle therapy. Barcelona, *ESTRO 37. Radiother. Oncol.* 127: S42-S43.

Contribution of Authors: Pedro Andreo provided useful guidance, discussions and provided help with the interpretation of ICRU recommended values and uncertainties. Arthur Lalonde provided help with initial MC simulations. Gary Royle provided input in form of discussion. Hugo Bouchard guided the development of the mathematical model, helped with the implementation and supervised the project. All other work was performed by myself.

4.1 Introduction

With DECT being on the edge to clinical implementation for radiotherapy, one major remaining source of uncertainty lies in the determination of the mean excitation energy, or I-value, of patient tissues. The portion of range uncertainties arising from I-values was previously estimated to 1.5% [9] and is currently taken into account as such in clinical treatment margins.

In practice, the I-values for compounds are calculated from elemental I-values using the Bragg additivity rule (BAR). Current clinically used elemental I-values were estimated by Berger and Seltzer in 1982 [34], and in 1984 those values were adapted as a recommendation in the ICRU report 37 [27], and later taken over for the use in proton and ion radiotherapy in ICRU report 49 [77]. Seltzer and Berger used a large set of different stopping power measurements of several compounds from previous publications [80, 81, 82, 83, 30, 31, 84, 85, 86, 87] and derived the compound I-values and their measurement uncertainties. Based on those compound measurements, they estimated two sets of elemental I-values to use in compounds in combination with the BAR, one set for gases and one set for liquids and solids. It is important to note that the I-value of an element or a molecule depends on whether it is unbound or bound, and the type of chemical bond. Hence, elemental I-values for the use in compounds are different from I-values of

unbound elements. Since the work of Berger and Seltzer was adapted to ICRU recommendations, their elemental I-values for the use in compounds are in clinical use.

The estimation of range uncertainties arising from I-values has always been challenging. Andreo [36] showed differences in beam depths of $0.3\,\frac{g}{cm^2}$ for a proton beam when the water I-value varies from 67 eV to 80 eV, covering the variety of values proposed in literature. Differences get larger when considering different tissue types or different particle species. Besemer *et al.* [88] performed a variation study that uniformly varies the tissue I-values, and evaluated the influence on patient dose distributions. They showed that a 10% variation of I-values influences the R_{80} beam range by up to 4.8 mm, and resulting dose distributions by up to 3.5%. Although the uncertainties on I-values can be relatively high, two studies by Yang *et al.* [51] and De Smet *et al.* [78] suggest that the resulting RSP values are much lower since correlations between water and medium need to be taken into account. Another study by Doolan *et al.* [89] investigated the influence of different correction terms to the Bethe formula on the calculated stopping power. They suggest using the I-value as a free parameter to optimize according to which corrections to the Bethe formula are used, in order to avoid systematic errors on RSP values. Further recent work suggests the estimation of a patient-specific tissue I-value from MRI imaging [90]. In this study, the I-value of tissue is parametrized as a function of three components: water content, mass fraction of organic tissues, and mass fraction of mineralized tissues. Those three quantities are either directly measurable with MRI (water content) or can be related to MR-measurable quantities (hydrogen content or hydroxyapatite).

The aim of this work is to revise the currently used elemental I-values. Since the work of Seltzer and Berger, several new stopping power measurements were performed [32, 91, 92, 33, 28] which can be included into the estimation of elemental I-values for the use in compounds. We develop a mathematical model to find, based on old and new measurement data, an optimal set of

elemental I-values for the used in compounds. Furthermore, our model establishes an uncertainty budget on our newly found set of elemental I-values as well as on the BAR. Our uncertainty budget allows the propagation of uncertainties from elemental I-values to RSPs of tissues and ultimately to beam ranges, to give a rigorous estimate of range uncertainties arising from I-values.

4.2 Materials and methods

4.2.1 Optimal elemental I-values to estimate compounds I-values

Let us define a series of M media indexed by $i = 1, \ldots, M$ consisting of N elements and with given elemental weights $w_{\text{med},ij}$, with $j = 1, \ldots, N$. The Bragg additivity rule (BAR) allows an estimation of the mean excitation energy $I_{\text{med},i}$ of the i-th medium using the weighted sum of the logarithmic elemental mean excitation energies:

$$\ln I_{\text{med},i} \approx \sum_{j=1}^{N} \lambda_{\text{med},ij} \ln I_{\text{med},j} \tag{4.1}$$

where $\lambda_{\text{med},ij}$ is the fraction of electrons from the j-th element in the i-th medium and given by

$$\lambda_{\text{med},ij} = \frac{w_{\text{med},ij} \frac{Z_j}{A_j}}{\left(\frac{Z}{A}\right)_{\text{med},i}} \tag{4.2}$$

with Z_j and A_j the atomic number and molar mass of the j-th element, and $\left(\frac{Z}{A}\right)_{\text{med},i}$ is the number of electrons per unit mass in the medium (in mol/g). Using matrix notation, the BAR can be written as the following estimator

$$\mathbf{y}_{\text{med}} \approx \hat{\mathbf{y}}_{\text{med}} \equiv \mathbf{\Lambda}_{\text{med}} \hat{\mathbf{y}}_{\text{elem}} \tag{4.3}$$

where $\mathbf{y}_{\text{med}}$ is a $M \times 1$-dimensional array containing the logarithm of experimental I-values and $\hat{\mathbf{y}}_{\text{elem}}$ is an array of dimension $N \times 1$ containing the optimized

logarithm of elemental I-values for use the BAR defined as

$$\hat{y}_{\text{elem},j} \equiv \ln \hat{I}_j. \tag{4.4}$$

The matrix $\mathbf{\Lambda}_{\text{med}}$ of dimension $M \times N$ contains the fractions of electrons for the respective materials and its elements are written as $\lambda_{\text{med},ij}$, corresponding tor the i-th medium and j-th element.

We propose to determine a new set of optimized elemental I-values, i.e., $\hat{I}_j$, by finding the weighted least square solution of equation 4.3. To take measurement and model uncertainties into account, we introduce weighting factors accounting for uncertainties:

$$\omega_i = \frac{1}{\sqrt{u^2_{\text{med},i}}} \tag{4.5}$$

with $u_{\text{med},i}$ being the relative uncertainty of the I-value measurement of the i-th medium. Note that because these uncertainties represent the absolute uncertainty of the natural logarithm of the I-value, they are reported in relative uncertainty on the I-value (i.e., in %). These weighting factors multiply individually both sides of the equation system 4.3 to account for uncertainties, leading to a new equation system

$$\tilde{\mathbf{y}}_{\text{med}} \approx \tilde{\mathbf{\Lambda}}_{\text{med}} \hat{\mathbf{y}}_{\text{elem}} \tag{4.6}$$

where the elements of $\tilde{\mathbf{y}}_{\text{med}}$ are $\omega_i y_{\text{med},i}$ and the elements of $\tilde{\mathbf{\Lambda}}_{\text{med}}$ are $\omega_i \lambda_{\text{med},ij}$. We can now find the least square solution to equation 4.6:

$$\begin{aligned} \widehat{\mathbf{y}}_{\text{elem}} &= \left(\tilde{\mathbf{\Lambda}}^{\text{T}}_{\text{med}} \tilde{\mathbf{\Lambda}}_{\text{med}}\right)^{-1} \tilde{\mathbf{\Lambda}}^{\text{T}}_{\text{med}} \tilde{\mathbf{y}}_{\text{med}} \\ &= \tilde{\mathbf{M}} \tilde{\mathbf{y}}_{\text{med}}, \end{aligned} \tag{4.7}$$

with $\widehat{\mathbf{y}}_{\text{elem}}$ being the estimation of the optimized logarithmic elemental I-values and $\tilde{\mathbf{M}}$ a projection matrix (from the measurement to the solution) defined to ease the notation. To find the uncertainties on elemental I-values, we construct

the covariance matrix of $\widehat{\mathbf{y}}_{\text{elem}}$ as follows:

$$\mathbf{V}\left(\widehat{\mathbf{y}}_{\text{elem}}\right) = \tilde{\mathbf{M}} V\left(\tilde{\mathbf{y}}_{\text{med}}\right) \tilde{\mathbf{M}}^{\text{T}} + u_{\text{BAR}}^2 \mathbb{1}_{N\times N}, \tag{4.8}$$

with $\mathbf{V}\left(\tilde{\mathbf{y}}_{\text{med}}\right)$ being the covariance matrix on measured material I-values with each line weighted by its corresponding ω_i. Note that because the measurement are assumed independent, $\mathbf{V}\left(\tilde{\mathbf{y}}_{\text{med}}\right)$ is diagonal. It should be kept in mind that the uncertainties u_{med} on the media are the measurement uncertainties quoted in the respective publication. It cannot be assumed that uncertainty estimations are consistent over the whole data set, hence these inconsistencies are reflected in the calculated weights.

Equation 4.8 is defined by combining two terms. The first one is obtained by applying the rule of uncertainty propagation on equation 4.7, since the Jacobian $(\partial\widehat{\mathbf{y}}_{\text{elem}}/\partial\tilde{\mathbf{y}}_{\text{med}})$ is the projection matrix $\tilde{\mathbf{M}}$. The second term is added to account for the model uncertainty. Indeed, the rule of uncertainty propagation can only yield accurate uncertainty estimations if the model is exact. Because the BAR is not completely accurate, it is judicious to add a model uncertainty component u_{BAR}^2 affecting each optimized values individually and in an independent manner. This way, we divide the uncertainties involved in the estimation of elemental I-values into experimental type A uncertainties and model-related type B uncertainties. The resulting $\mathbf{V}\left(\widehat{\mathbf{y}}_{\text{elem}}\right)$ is a non-diagonal square matrix of dimensions $N \times N$ accounting for statistical correlations in the solution.

The solution expressed in equation 4.7 using experimental data yields a new set of elemental I-values $\hat{I}_j$ for use with the BAR. The experimental data were taken from different sources as listed in table 4.1. They include the data provided in ICRU report 37, table 5.3 of [93] and more recent publications. Like Seltzer and Berger, we divide the data into two groups: 1) gases; 2) liquids and solids. For gases, the data used in this work are the same than the ones used by Seltzer and Berger [34] to determine the recommended elemental I-values in ICRU report 37. For liquids and solids, we added data published in recent

literature (see table 4.1, numbers 11-15). In total, we use 74 liquids and solids for calibration, including six different I-values for water [28, 29, 30, 31, 32, 33]. We obtain two sets of elemental I-values which are optimized for the use in compounds (one set for gases, one for liquids and solids) in combination with the BAR. We use equation 4.8 to report uncertainty values u_{elem} for the newly determined I_{elem}. The here reported uncertainties are standard uncertainties (68% confidence interval). Some of the data used for our analysis, especially the ones used by Seltzer and Berger, are quoted for a 90% confidence interval, as reported in ICRU report 37, footnote 10. Whenever this was the case, the uncertainties were divided by a factor of 1.6 to convert from the 90% to the 68% confidence interval.

To test the validity of our data, we perform a self-consistency test. For this, we use the optimized elemental I-values to reproduce the calibration data set. The root mean square (RMS) errors between actual and predicted calibration data are compared with predicted data using ICRU-recommended elemental I-values. Again, we separate gases from liquids and solids.

4.2.2 Estimation of $u_{\mathbf{BAR}}$

While $u_{\text{med},i}$ can be derived from experimental uncertainties, u_{BAR} needs to be estimated using a model. To estimate u_{BAR}, we use the calibration data set as listed in table 4.1, and calculate the residual error $\mathbf{r}$ between the estimated and experimentally measured values for $\mathbf{y}_{\text{med}}$:

$$\begin{aligned}\mathbf{r} &= \hat{\mathbf{y}}_{\text{med}} - \mathbf{y}_{\text{med}} \\ &= \mathbf{\Lambda}\left(\tilde{\mathbf{\Lambda}}_{\text{med}}^{\text{T}}\tilde{\mathbf{\Lambda}}_{\text{med}}\right)^{-1}\tilde{\mathbf{\Lambda}}_{\text{med}}^{\text{T}}\tilde{\mathbf{y}}_{\text{med}} - \mathbf{y}_{\text{med}} \\ &= \tilde{\mathbf{K}}\tilde{\mathbf{y}}_{\text{med}} - \mathbf{y}_{\text{med}} \\ &= \left(\mathbf{K} - \mathbb{1}_{N\times N}\right)\mathbf{y}_{\text{med}}.\end{aligned} \tag{4.9}$$

with $\tilde{\mathbf{K}} = \mathbf{\Lambda}\tilde{\mathbf{M}}$ and where the elements of $\mathbf{K}$ equal the ones of $\tilde{\mathbf{K}}$ divided by the weighting factors, i.e., $\mathrm{K}_{ij} = \frac{1}{w_i}\tilde{\mathrm{K}}_{ij}$. We can now estimate the covariance matrix $\mathbf{V}(\mathbf{r})$ as

Table 4.1: Literature reporting on the I-values of compounds, by either dipole-oscillator strength distributions or measurements of the energy loss. Numbers 1-10 were used by Seltzer and Berger (1981) to assign the elemental I-values for the use in compounds with the BAR.

Number	Source	Year	Method	Elements involved
1	Zeiss *et al.*	1977	dipole oscillator strength	N, H, O
2	Jhanwar *et al.*	1981	dipole oscillator strength	H, C
3	Bichsel and Hilko	1981	α-particle beam	C, O
4	Thomas and Meath	1977	dipole oscillator strength	H, C
5	Thompson	1952	proton beam	H, C, N, O, Cl
6	Nordin and Henkelmann	1979	pion beam	H, O
7	Bader *et al.*	1956	low energy proton beam	F, Ca
8	Painter *et al.*	1980	dielectric-response functions	H, C
9	Tschalär and Bichsel	1968	proton beam	H, C, O, Si
10	Ashley	1979	dielectric-response functions	H, C
11	Bichsel and Hiraoka	1992	proton beam	H, O
12	Hiraoka *et al.*	1993	proton beam	H, C, N, O, F, Cl
13	Hiraoka *et al.*	1994	proton beam	H, C, N, O, F, Cl, Si, P, Ca
14	Bichsel *et al.*	2000	carbon beam	H, O
15	Kumazaki *et al.*	2007	proton beam	H, O

$$\mathbf{V}(\mathbf{r}) = (\mathbf{K} - \mathbb{1}_{N\times N})\, V(\mathbf{y}_{\text{med}})\, (\mathbf{K} - \mathbb{1}_{N\times N})^{\text{T}} + u^2_{\text{BAR}}\mathbb{1}_{N\times N}. \tag{4.10}$$

To solve for u^2_{BAR}, we find the value such that the sum of the residuals squared normalized to their variance equals its number of degrees of freedom, that is:

$$\sum_{i=1}^{N} \frac{r_i^2}{\text{V}_{ii}(r)} = N - M. \tag{4.11}$$

Note that the approach is based on the assumption that the experimental data for a particular element follows a Gaussian distribution. This results in the sum of above following a chi-square distribution with expectation value equaling its number of degrees of freedom $N - M$, N being the number of experimental data and M the number of optimized elemental I-values.

4.2.3 Application of optimal elemental I-values to water and reference human tissues

Using the optimized set of elemental I-values, we determine compound I-values for water and a set of 70 human reference tissues [48, 49]. The compound I-values are compared to the results obtained with ICRU 37 recommended values. We establish the uncertainties on compound I-values using the covariance matrix of the elemental I-values $\mathbf{V}\left(\widehat{y}_{\text{elem}}\right)$:

$$\mathbf{V}\left(\widehat{y}_{\text{med}}\right) = \Lambda_{\text{med}} V\left(\widehat{y}_{\text{elem}}\right) \Lambda^{\text{T}}_{\text{med}}. \tag{4.12}$$

4.2.4 Uncertainties on RSPs

Once the uncertainties on the mean excitation energies are determined, it is possible to propagate these into stopping power uncertainties. In this way, we quantify the uncertainty on the RSP of medium to water originating from uncertainties on I-values. Since the uncertainties of medium and water can be correlated depending on the water content of the medium [78, 51], it is important to consider covariances. We can formulate the RSP in terms of

relative electron density ρ_{e} and stopping number as

$$S_{\mathrm{med}} = \rho_{\mathrm{e}} \frac{L_{\mathrm{med}}}{L_{\mathrm{w}}}. \tag{4.13}$$

where L_{med} and L_{w} are the stopping numbers of the medium and water, respectively. Using Bethe's model and the BAR, the stopping number of an arbitrary medium is expressed as

$$L = \ln\left(\frac{2m_{\mathrm{e}}c^2\beta^2}{1-\beta^2}\right) - \beta^2 - \sum_{k=1}^{N} \lambda_k y_k \tag{4.14}$$

with m_{e} the electron mass, c the speed of light and β the particle velocity relative to c. The derivative of the stopping number with respect to $\ln I_i$ is then found to be

$$\frac{\partial L}{\partial y_i} = -\lambda_i. \tag{4.15}$$

We can now express the derivative of the RSP with respect to $\ln I_i$ as

$$\begin{aligned} \frac{\partial S}{\partial y_i} &= \rho_{\mathrm{e}} \frac{\partial \frac{L_{\mathrm{med}}}{L_{\mathrm{w}}}}{\partial y_i} \\ &= \rho_{\mathrm{e}} \frac{L_{\mathrm{med}}\lambda_{\mathrm{w},i} - L_{\mathrm{w}}\lambda_{\mathrm{med},i}}{L_{\mathrm{w}}^2}. \end{aligned} \tag{4.16}$$

The variance on the RSP can now be written as using the following rule

$$\begin{aligned} V(S) &= \left(\frac{\partial S}{\partial \mathbf{y}}\right)^{\mathrm{T}} V(\mathbf{y}) \left(\frac{\partial S}{\partial \mathbf{y}}\right) \\ &= \begin{pmatrix} \frac{\partial S}{\partial y_1} \\ \vdots \\ \frac{\partial S}{\partial y_N} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \mathrm{COVAR}(y_1, y_1) & & \mathrm{COVAR}(y_1, y_N) \\ \vdots & \ddots & \vdots \\ \mathrm{COVAR}(y_N, y_1) & \cdots & \mathrm{COVAR}(y_N, y_N) \end{pmatrix} \begin{pmatrix} \frac{\partial S}{\partial y_1} \\ \vdots \\ \frac{\partial S}{\partial y_N} \end{pmatrix}. \end{aligned} \tag{4.17}$$

4.2.5 Uncertainties on beam ranges

To quantify the impact on beam ranges, we perform MC transport simulations of a pristine proton beam in homogeneous media (volume: $30\times30\times30\,\mathrm{cm}^3$).

We score the energy loss and position of each interaction of the beam with the medium. We choose water and five different human reference tissues (Adipose 3, skeletal muscle 1, brain white matter, femur whole, cortical bone) relevant to proton therapy. For every material, four simulations are performed: 1) using ICRU-recommended I-values; 2) using our suggested I-values; 3) using the upper uncertainty limit and 4) using the lower uncertainty limit. For the simulations, we use the Geant4 code (Version 10.03.p02) with the QBBC physics package [94]. We simulate proton beams of 173 MeV using 10^6 particles per beam and a 1 mm cut-off value for secondary particles.

4.3 Results

4.3.1 Optimal elemental I-values to estimate compounds I-values

The proposed approach results in a set of optimized elemental I-values for the use with the BAR with compounds. Using the same measured compound I-values than Seltzer and Berger and more recent literature on measured I-values, we calculate optimized elemental I-values for the use in gases and for the use in liquids and solids separately. Our optimized elemental I-values differ from the ones suggested by Seltzer and Berger, as tabulated in tables 4.2 and 4.3. Table 4.4 shows the correlation coefficients of the optimized elemental I-values for liquids and solids. We use both sets of elemental I-values to perform a self-consistency test on the calibration data. Using the ICRU 37 recommended elemental I-values suggested by Seltzer and Berger, we observe RMS errors of 1.02% (gases) and 6.17% (liquids and solids) when using the BAR to predict the underlying experimental I-values. Using our optimized elemental I-values, this prediction error can be reduced to 0.05% (gases) and 5.19% (liquids and solids). The model uncertainty arising from the BAR, i.e., u_{BAR}, is quantified as 4.42%.

Table 4.2: Comparison of elemental I-values (in eV) for the use in gas compounds. The values recommended by Seltzer and Berger (ICRU 37) are compared to the values determined with our proposed assignment scheme.

Gases			
Element	Seltzer and Berger	This work	Uncertainty [eV]
H	19.20	21.54	0.74
C	70.00	66.75	1.08
N	82.00	79.59	1.15
O	97.00	95.17	1.01

Table 4.3: Comparison of elemental I-values in eV for the use in liquid and solid compounds. The values recommended by Seltzer and Berger (ICRU 37) are compared to the values determined with our proposed assignment scheme.

Liquids and solids			
Element	Seltzer and Berger	This work	Uncertainty [eV]
H	19.20	22.07	1.32
C	81.00	79.91	3.61
N	82.00	77.91	3.86
O	106.00	107.44	4.88
F	112.00	136.24	6.28
Al	187.58	191.69	11.13
Si	195.50	150.47	7.60
P	195.50	199.39	42.45
Cl	180.00	175.13	7.91
Ca	215.80	258.11	16.61

4.3.2 Application of optimal elemental I-values to water and reference human tissues

Using our method, we estimate the compound I-value of water to 78.73±2.89 eV. This value is in good agreement with the recent recommendation given in ICRU 90 [95], which is based on the value 78±2 eV given in Andreo *et al.* [96]. The compound I-values of 70 reference human tissues are listed in tables 4.5 and 4.6. The uncertainty of the values suggested herein and the difference with Seltzer and Berger recommended values are also listed. In figure 4.1, we show the resulting uncertainties on tissue I-values when covariances are not taken into account. With our model, we obtain uncertainties between $\min\{u_{\text{tissue}}\} = 1.82\,\text{eV}$ (mammary gland) and $\max\{u_{\text{tissue}}\} = 3.38\,\text{eV}$ (cortical bone). If statistical correlations between optimized elemental I-values are neglected, these values increase to $\min\{u_{\text{tissue}}\} = 1.92\,\text{eV}$ (mammary gland)

Table 4.4: Correlation coefficients of the uncertainties of elemental I-values.

Element	H	C	N	O	F	Al	Si	P	Cl	Ca
H	1.00									
C	-0.12	1.00								
N	-0.02	-0.00	1.00							
O	-0.12	0.03	0.01	1.00						
F	0.04	-0.02	0.00	-0.01	1.00					
Al	0.09	-0.02	-0.01	-0.04	0.01	1.00				
Si	0.03	-0.01	-0.00	-0.01	0.00	0.01	1.00			
P	0.06	-0.02	-0.00	-0.02	0.03	0.02	0.01	1.00		
Cl	0.01	-0.01	0.00	-0.00	-0.00	0.00	0.00	0.00	1.00	
Ca	0.04	-0.02	-0.01	-0.02	-0.02	0.01	0.00	-0.35	0.01	1.00

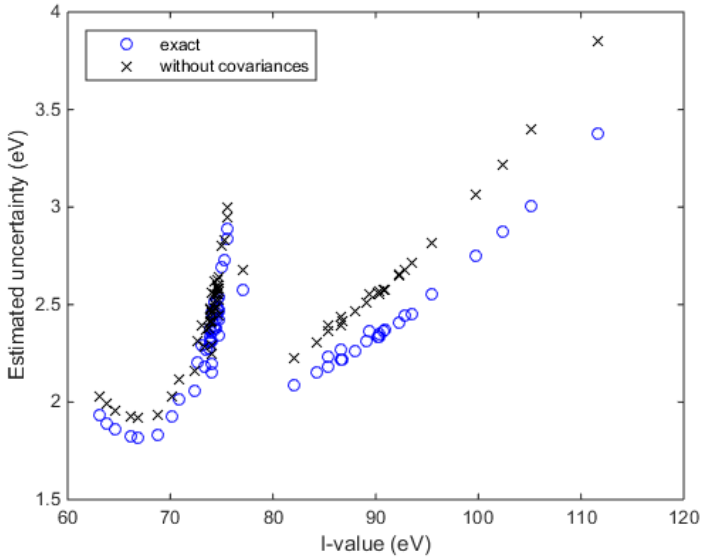


Figure 4.1: Calculated uncertainties on compound I-values for 70 human reference tissues. This graph also shows that neglecting statistical correlations between optimized elemental I-values leads to an overestimation of the uncertainties in compounds.

and $\max\{u_{\text{tissue}}\} = 3.85\,\text{eV}$ (cortical bone).

4.3.3 Uncertainties on RSPs

We use our optimized I-values to calculate RSP values for 70 human reference tissues. Figure 4.2 shows the uncertainties on RSP values of 70 human reference tissues, arising from uncertainties on I-values only. We observe uncertainties on RSP values between 0.002% (mammary gland) and 0.44% (adipose tissue 3). The uncertainties observed are the smallest for the soft tissues since their water content is the highest between adipose tissue, soft tissues,

lcccc

Table 4.5: Compound I-values determined with the elemental I-values recommended by Seltzer and Berger (ICRU 37) and compared to those suggested in this work. All I-values and standard uncertainties are given in eV.

Tissue	Seltzer and Berger	This work	Uncertainty	I-value difference
Adipose tissue 1	66.20	68.28	1.82 (2.67%)	2.09 (3.15%)
Adipose tissue 2	64.66	66.62	1.86 (2.79%)	1.96 (3.04%)
Adipose tissue 3	63.12	64.95	1.93 (2.98%)	1.84 (2.91%)
Adrenal gland	70.83	73.33	2.01 (2.74%)	2.50 (3.52%)
Aorta	74.78	77.50	2.34 (3.02%)	2.72 (3.64%)
Blood whole	74.78	77.62	2.47 (3.18%)	2.84 (3.79%)
Brain cerebrospinal fluid	75.52	78.93	2.89 (3.66%)	3.40 (4.51%)
Brain gray matter	74.34	77.36	2.52 (3.25%)	3.02 (4.06%)
Brain white matter	72.67	75.39	2.21 (2.92%)	2.72 (3.74%)
C4 including cartilage male	89.41	93.34	2.37 (2.53%)	3.93 (4.39%)
Cartilage	77.14	80.04	2.58 (3.22%)	2.90 (3.76%)
Clavicle scapula	92.25	96.53	2.40 (2.49%)	4.27 (4.63%)
Connective tissue	73.97	76.30	2.15 (2.82%)	2.34 (3.16%)
Cortical bone	111.63	117.81	3.38 (2.87%)	6.17 (5.53%)
Cranium	99.69	104.65	2.75 (2.63%)	4.97 (4.98%)
D6L3 including cartilage male	85.44	89.01	2.23 (2.51%)	3.58 (4.19%)
Eye lens	74.03	76.45	2.20 (2.87%)	2.42 (3.27%)
Femur Humerus spherical head	85.43	89.09	2.18 (2.45%)	3.66 (4.29%)
Femur conical trochanter	86.69	90.47	2.22 (2.45%)	3.78 (4.36%)
Femur cylindrical shaft	105.13	110.64	3.01 (2.72%)	5.51 (5.24%)
Femur total bone	90.24	94.31	2.33 (2.47%)	4.08 (4.52%)
Femur whole specimen	90.34	94.44	2.34 (2.47%)	4.10 (4.54%)
Gallbladder bile	75.03	78.23	2.69 (3.44%)	3.20 (4.26%)
Heart 1	73.43	76.13	2.27 (2.98%)	2.70 (3.67%)
Heart 2	73.91	76.72	2.37 (3.09%)	2.81 (3.80%)
Heart 3	74.61	77.52	2.49 (3.21%)	2.91 (3.91%)
Heart blood-filled	74.39	77.21	2.42 (3.14%)	2.83 (3.80%)
Humerus cylindrical shaft	93.56	97.96	2.45 (2.50%)	4.40 (4.70%)
Humerus total bone	92.23	96.50	2.40 (2.49%)	4.27 (4.63%)
Humerus whole specimen	88.06	91.98	2.26 (2.46%)	3.92 (4.45%)
Innominate female	92.82	97.08	2.45 (2.52%)	4.27 (4.60%)
Innominate male	90.75	94.80	2.37 (2.50%)	4.05 (4.46%)
Kidney 1	73.90	76.62	2.31 (3.01%)	2.72 (3.68%)
Kidney 2	74.28	77.10	2.40 (3.11%)	2.82 (3.80%)
Kidney 3	74.60	77.52	2.48 (3.20%)	2.92 (3.91%)
Liver 1	73.84	76.60	2.33 (3.04%)	2.77 (3.75%)
Liver 2	74.30	77.08	2.38 (3.09%)	2.78 (3.75%)
Liver 3	74.69	77.48	2.42 (3.12%)	2.79 (3.73%)
Lung deflated	74.72	77.59	2.48 (3.19%)	2.88 (3.85%)
Lymph	75.23	78.43	2.73 (3.48%)	3.20 (4.25%)
Mammary gland 1	66.79	68.81	1.82 (2.64%)	2.03 (3.04%)
Mammary gland 2	70.11	72.46	1.93 (2.66%)	2.35 (3.35%)
Mammary gland 3	73.84	76.53	2.31 (3.02%)	2.69 (3.65%)
Mandible	102.35	107.56	2.88 (2.67%)	5.21 (5.09%)
Muscle skeletal 1	73.74	76.40	2.28 (2.98%)	2.66 (3.60%)
Muscle skeletal 2	74.08	76.84	2.36 (3.07%)	2.75 (3.71%)
Muscle skeletal 3	74.72	77.58	2.47 (3.18%)	2.85 (3.82%)
Ovary	74.63	77.60	2.52 (3.25%)	2.97 (3.98%)

Table 4.6: Continued: Comparison of ICRU 37 recommended I-values with our suggested I-values for human reference tissues.

Tissue	Seltzer and Berger	This work	Uncertainty	I-value difference
Pancreas	73.11	75.90	2.29 (3.02%)	2.79 (3.82%)
Prostate	74.70	77.67	2.54 (3.27%)	2.97 (3.97%)
Red marrow	68.81	70.94	1.83 (2.59%)	2.13 (3.10%)
Ribs 10th	95.48	100.00	2.55 (2.55%)	4.53 (4.74%)
Ribs 2nd 6th	90.33	94.37	2.35 (2.49%)	4.04 (4.47%)
Sacrum female	89.16	93.10	2.31 (2.48%)	3.93 (4.41%)
Sacrum male	84.25	87.71	2.15 (2.45%)	3.46 (4.11%)
Skin 1	72.43	74.82	2.06 (2.75%)	2.39 (3.30%)
Skin 2	73.35	75.87	2.18 (2.88%)	2.52 (3.44%)
Skin 3	74.06	76.74	2.31 (3.02%)	2.68 (3.62%)
Small intestine wall	74.09	77.02	2.46 (3.19%)	2.93 (3.95%)
Spleen	74.58	77.43	2.45 (3.16%)	2.85 (3.82%)
Sternum	82.03	85.28	2.09 (2.45%)	3.25 (3.96%)
Stomach	73.87	76.67	2.37 (3.10%)	2.81 (3.80%)
Testis	74.35	77.33	2.51 (3.25%)	2.98 (4.01%)
Thyroid	74.35	77.23	2.45 (3.18%)	2.88 (3.88%)
Trachea	74.44	77.20	2.38 (3.08%)	2.76 (3.71%)
Urine	75.55	78.88	2.84 (3.60%)	3.33 (4.40%)
Vertebral column C4 excluding cartilage	90.83	94.90	2.37 (2.49%)	4.07 (4.48%)
Vertebral column D6L3 excluding cartilage	86.59	90.27	2.22 (2.46%)	3.67 (4.24%)
Vertebral column whole	86.65	90.34	2.27 (2.51%)	3.70 (4.27%)
Water	75.31	78.73	2.89 (3.67%)	3.41 (4.53%)
Yellow marrow	63.78	65.66	1.89 (2.88%)	1.88 (2.95%)

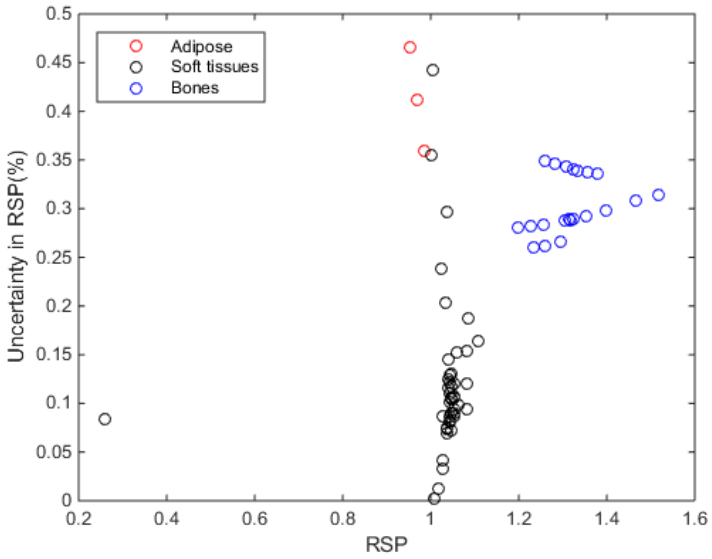


Figure 4.2: Calculated uncertainties on RSPs for 70 human reference tissues, arising from the uncertainties on compound I-values.

and bones.

4.3.4 Uncertainties on beam ranges

Figure 4.3 shows the percentage depth dose (PDD) curves for water and five human reference tissues, each using four different sets of I-values: 1) the ICRU-

recommended values; 2) the optimized I-values resulting from this work, 3) the optimized I-values resulting from this work plus 1 standard deviation and 4) the optimized I-values resulting from this work minus 1 standard deviation. We observe differences in the range of the distal 80% of the maximum dose (R_{80}) between ICRU-recommended values and our values of 0.75 mm (adipose tissue 3) - 1.10 mm (water). We find range uncertainties between 0.31% and 0.47%, with the lowest uncertainty found in femur tissue, while the highest uncertainty is found in water (see table 4.7).

Table 4.7: Calculated beam ranges in terms of R_{80} (in mm) using MC proton beam transport simulations. The uncertainties reported are resulting from the uncertainties on our optimized I-values and the differences are taken between ranges simulated with ICRU-recommended I-values and ranges simulated with our optimized I-values.

	Liquids and solids			
Material	Range ICRU	Range this work	Uncertainty [%]	Difference [%]
Water	202.25	203.35	0.95 (0.47%)	1.10 (0.54%)
Adipose tissue 3	211.97	212.72	0.79 (0.37%)	0.75 (0.35%)
Muscle skeletal 1	194.01	194.86	0.72 (0.37%)	0.85 (0.44%)
Brain white matter	194.68	195.55	0.72 (0.37%)	0.88 (0.45%)
Femur whole	151.49	152.34	0.47 (0.31%)	0.85 (0.56%)
Cortical bone	119.46	120.28	0.50 (0.42%)	0.82 (0.68%)

4.4 Discussion

In this work, we investigate RSP and range uncertainties arising from mean excitation energies. We establish a mathematical model to optimize elemental I-values for the use in gases and liquids and solids with the BAR. To calculate our optimized I-values and establish an uncertainty budget, we utilize I-value and stopping power measurements from literature, most of which were used by Seltzer and Berger to establish the ICRU 37 recommended values, however, we also include more recent measurements. To test consistency with Seltzer and Bergers publication, we applied our methods to their original dataset. The results obtained with our method differ from the elemental I-values proposed by Seltzer and Berger, indicating that there is a methodological difference between the two methods. The set of optimized elemental I-values for the

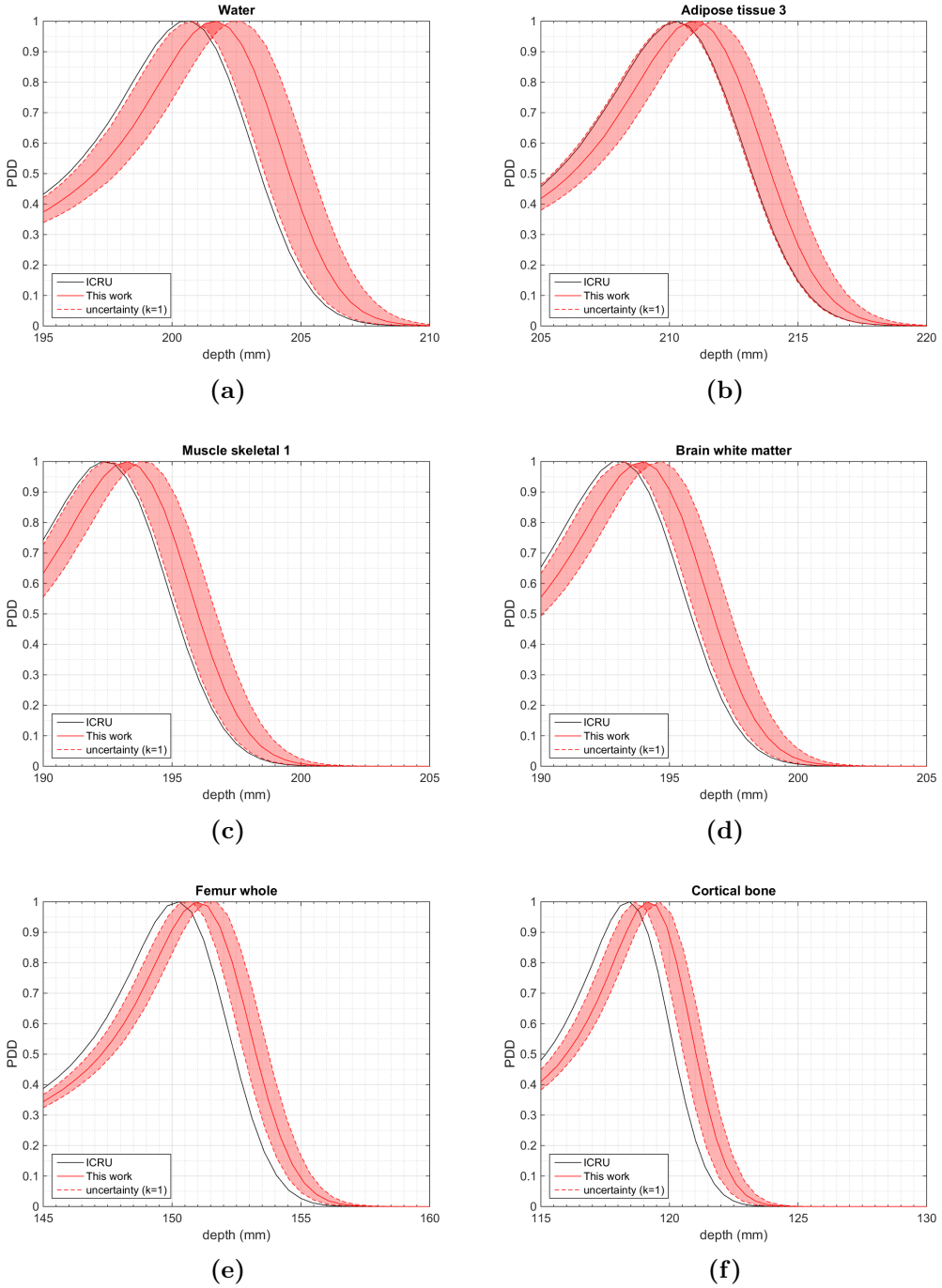


Figure 4.3: Percentage depth dose curves of a pristine proton beam with an initial energy of 173 MeV, simulated in homogenous media: a) Water, b) adipose tissue 3, c) muscle skeletal 1, d) brain white matter, e) femur whole and f) cortical bone. Shown are curves using the ICRU-recommended I-values (black line), our suggested I-values (red solid line) and uncertainty limits (dashed lines).

use with the BAR in compounds and the reported uncertainties that can be used to accurately assess the uncertainties on tissue I-values. Furthermore, we provide an estimation of the uncertainty coming from the BAR itself and include this uncertainty in our model. Our model allows the propagation of uncertainties to RSP values and beam ranges, providing a better understanding of the resulting uncertainties in proton therapy treatment planning.

Elemental I-values of liquids and solids are of special interest for proton therapy treatment planning. We propose a set of values which differs from the values recommended in ICRU 37. This is due to the availability of more recent data and the fact that the underlying model used to find the optimized I-values differs from the methods used by Seltzer and Berger to determine the originally proposed values. For the majority of our proposed values (C, N, O, Al, P, Cl), the ICRU-recommended values are within the uncertainty budget given by our model. However, few exceptions are observed. We find considerably higher elemental I-values for the elements H, F, and Ca than Seltzer and Berger. In turn, our values for Si is lower. While the optimized value for P is close to the ICRU recommended value (195.5 eV versus 199.93 eV), we observe a high uncertainty of 42.45 eV. The high uncertainty can be explained due to the lack of data available for P. From the 74 liquids and solids, only 3 materials contain traces of P. The high observed uncertainty, however, is of little concern to clinical applications. First of all, P is only abundant as a trace element in the human body. Secondly, it is only observed in bones, where we observe a strong statistical anti-correlation with Ca, compensating for the high uncertainty in P.

The calibration materials used herein are taken from various sources of literature, most of them were already utilized by Seltzer and Berger to recommend the currently clinically applied elemental I-values. As a self-consistency study, we evaluate the accuracy of our optimized elemental I-values in comparison to ICRU-recommended elemental I-values to predict the I-values of

the calibration material. Our optimized elemental I-values show a lower RMS prediction error than the ICRU-recommended values, indicating that our values are well suited to predict the I-values of tissues when the BAR is used. When calculating the I-values of 70 human reference tissues, we observe generally larger tissue I-values using the optimized elemental I-values. Using our technique and optimized elemental I-values, we achieve an I-value for water of 78.73 ± 2.89 eV, which is in good agreement with the value recommended in the recently published ICRU 90 report. Using the elemental I-values recommended in ICRU 37, the I-value for water is estimated as 75.32 eV. Please note that the value quoted in ICRU 90 is derived from water I-values only. In this work, we include water but also various materials to derive elemental I-values which are then used to calculated the I-value of water. An alternative approach could be to fix the I-value of water to the latest ICRU recommendation to obtain a set which is in better consistency with the ICRU report.

Our mathematical model incorporates estimates for uncertainty values associated with each optimized elemental I-value and allows the propagation of uncertainties on tissue I-value uncertainties. We show the importance of taking into account statistical correlations between uncertainties of the different elemental I-values. Our analysis shows that if statistical correlations are ignored, the uncertainties might be overestimated by up to 0.5%.

The method proposed herein allows the estimation of uncertainties on tissue RSPs resulting from I-value uncertainties. We observe the highest RSP uncertainties in adipose tissues and bones, which is expected as those are the tissues with a low water content. Those findings were already discussed in previous studies by Yang *et al.* [51] and De Smet *et al.* [78]. We observe low RSP uncertainties in soft tissues, which can be attributed to their high water content.

The resulting beam range uncertainties are assessed in water and five selected human reference tissues. We choose the tissues according to their importance to radiotherapy and abundance in the human body. Observed range uncertainties are between 0.31% and 0.47%for a 173 MeV proton beam. Overall, the ranges predicted using our optimized elemental I-values are systematically larger than the ranges calculated based on ICRU-recommended I-values. This is expected since our optimized I-values calculate larger tissue I-values than the ICRU recommendation.

It should be noted that the systematic shift in depth dose curves observed in the MC simulations will not necessarily be observed with more simplistic dose calculation algorithms as implemented in treatment planning systems since they usually use the stopping power relative to water for dose estimation. Additionally, it must be considered that MC based treatment planning systems are usually commissioned based on measured ranges. The incident energy and the I-value of water may be tuned such that the model matches the measured ranges. However, it should be emphasized that our model yields a more realistic I-value for water (and closer to the recent ICRU 90 recommendation) than ICRU 37. Future work will focus on the application of optimized I-values to CT scans of tissues, which can potentially prove the validity and superiority of one set of elemental I-values over the other.

4.5 Conclusion

We propose a new set of optimized elemental I-values for the use with the Bragg additivity rule in compounds. Our mathematical model establishes an uncertainty budget on elemental I-values that can be propagated to compound I-values, accounting for experimental uncertainties as well as the uncertainty on the Bragg additivity rule itself. With our model, we provide realistic uncertainties estimations on proton RSP and beam range values in human tissues. In conclusion, this chapter shows that the currently assumed range uncertainty originating from I-values as quoted in chapter 1 (1.5%) may be overestimated.

The here proposed elemental I-values and corresponding uncertainty budgets provide ground for a reassessment of those values and a possible reduction of range uncertainties by up to 1%.

Chapter 5

Final remarks

Radiotherapy with protons and ions is becoming widely available as an option for cancer treatment. A crucial part of treatment planning is the prediction of the range that the particles travel in the patient. In its current form, however, proton and ion therapy is not used to its full potential. Clinically applied safety margins were defined as early as 1985 [97], and are still in use in many centers nowadays. The use of large margins implies that a large amount of healthy tissues is irradiated. A reduction of range margins, however, needs to be performed in a careful manner and must be based on plenty of scientific evidence. With proton therapy becoming widely available, there is a clear need to carefully improve treatment accuracy by addressing the various sources of range uncertainties.

One approach to improve treatment accuracy is by improving tissue characterization based on CT scans. While DECT is widely available and in routine clinical use for radiological applications, SECT is still the preferred method for CT to RSP conversion in treatment planning. DECT, despite being quoted in literature to improve range predictions, has not yet found its way into clinical application due to the lack of data showing its safety and advantages in clinical situations.

An uncertainty closely connected to those from CT imaging and conversion to tissue is the uncertainty on the mean excitation energies. With DECT having shown a potential to improve tissue characterization, the question about

the contribution from I-values arises. The value currently applied to clinically used uncertainty budgets and range margins is based on an estimate and shows room for improvement.

A further source of uncertainty is the choice of dose calculation algorithms. It was demonstrated in literature that a simple ray tracing algorithm is not sufficient to predict the range in heterogeneous media and gives rise to large range uncertainties. To improve treatment, MC is needed to accurately model beam transport in patients.

The work described in this thesis has contributed to the field by addressing the above-mentioned issues. The clinical applicability of DECT was demonstrated by (i) showing the potential of DECT in the presence of imaging artifacts and by (ii) validating DECT predicted RSP values in an experimental setup with animal tissues. The uncertainties arising from lateral inhomogeneities were investigated (iii) using animal tissue measurements. The uncertainties arising from mean excitation energies were addressed by (iv) providing a new set of elemental I-values and a thorough uncertainty budget.

The potential of DECT and possible improvements compared to SECT were investigated in a theoretical study. This study suggests that DECT can improve the RSP predictions and reduce range uncertainties. In this study, different formalisms published in literature to use DECT to convert CT numbers into RSP values were compared and shown to perform differently well. The outcomes of this study might aid clinical scientists in their decision making on which method to implement clinically. It was the first study to compare methods suggested in literature and apply them to realistic DECT images. An important outcome of this study is the performance of DECT in the presence of image artifacts. Most importantly, it was shown that DECT is more sensitive to image noise than SECT, which can out-weight the benefits of DECT over SECT. Following this finding, a Bayesian approach to solve

RSPs from noisy DECT data was developed.

Following the theoretical study that pointed at a potential improvement of range predictions using DECT, an experimental validation of these findings was performed. In this study, the WER of fresh animal tissue samples was measured in a proton beam to provide a ground truth. The WER was then predicted from CT scans using both, SECT and DECT, and the WER prediction accuracy of both imaging modalities was compared. Results of this study show that a commonly used SECT method performs well, however DECT can improve tissue characterization for proton radiotherapy and thus reduce range uncertainties. The acquired data provide evidence for the clinical applicability of DECT and quote an uncertainty budget to estimate range margins, laying the ground for a clinical implementation of DECT for proton therapy planning. Future work will be translational and focus on the comparison of patient treatment plans and dose differences between SECT and DECT.

The uncertainties arising from lateral inhomogeneities and the choice of the dose calculation algorithm have been discussed in the past few years. It is generally accepted that MC needs to be implemented for treatment planning, and several research groups and vendors are working on the implementation and realization. In this thesis, an initial experimental investigation of the involved range errors with a known ground truth was performed to demonstrate the gain in accuracy when MC is chosen over ray tracing. Future work can focus on expanding and refining this study. Work has been done to improve the performance of a pencil beam algorithm in the presence of lateral inhomogeneities, which can be included in the present study. Additional work will focus on providing better MC input parameters.

The mean excitation energy is subject to discussion in recent literature. The aim of the study presented in this thesis was to improve knowledge and

understanding of uncertainties coming from I-values. Since I-values of a material cannot be known or measured, it is crucial to estimate a set of elemental I-values that can be used to best describe average I-values of human tissues. Hence a new set of I-values is proposed in this thesis, together with a complete uncertainty budget that allows the estimation of range uncertainties coming from I-values alone. The presented results confirm that the currently used uncertainty value of 1.5% is high, as could already be expected from the experimental validation of DECT predicted ranges. A necessary next step would be to validate the newly found elemental I-values in an experimental study and compare the results to currently used ICRU recommended elemental I-values.

Table 1.1 shows the current estimates of sources of range uncertainties in proton therapy. In this thesis, several of these numbers were challenged. The here presented work increases knowledge of several components of range uncertainties. The main achievements of this thesis were the validation of DECT for range prediction and the quantification of the associated range uncertainties. It was shown that DECT can improve range uncertainties as opposed to SECT in soft tissues and bones, and an estimate of the associated uncertainties is given. The current value of 0.5% should be adjusted to 0.19% for soft tissues and 1.05% for bones. Furthermore, the uncertainties from mean excitation energies were quantified in a rigorous study, giving a better estimate of resulting range uncertainties. According to the here presented data, the value can be reduced from 1.5% to 0.45% in soft tissues and 0.35% in bones. Table 5.1 shows the imaging-related sources of range uncertainty as quoted by Paganetti and our revised values. It can be concluded that the contribution of the imaging-related relative range uncertainties should be adjusted from the current value of 2.4% to 0.9% in soft tissues and 1.7% in bones.

One aspect of DECT that was not discussed in this thesis but will be addressed as future work is the ability to estimate nuclear interaction cross

Table 5.1: Imaging- related sources of range uncertainties as investigated in this thesis. The first column shows the currently assumed values listed in [9]. The second and third columns show the values derived within the scope of this thesis. All values for the range uncertainties are given relative to the total range and are based on 1.5 standard deviations. The bottom lines are the total uncertainties. Those uncertainties printed in bold were subject of evaluation in this thesis.

Source of range uncertainty	Paganetti 2012	This work: Soft tissues	This work: Bones
CT imaging and calibration	0.5%	0.19%	1.05%
CT conversion to tissue	0.2%		
CT grid size	0.3%	0.3%	0.3%
Mean excitation energies in tissues	1.5%	0.45%	0.35%
Total	2.4%	0.9%	1.7%

sections. Nuclear interaction cross sections of materials are needed for MC treatment planning. Currently, they are either retrieved from tabulated data after tissue segmentation or found via a calibration from SECT-determined CT numbers [98]. With DECT, it is possible to estimate the material compositions in form of elemental mass fractions. These can be used to calculate material-specific nuclear interaction cross sections.

Another aspect of DECT that can be of interest to future research is the ability to estimate lateral scattering of a pencil beam. Modern proton therapy techniques such as pencil beam scanning require accurate dose calculation techniques. Current pencil beam algorithms are fast but lack in accuracy when tissue inhomogeneities are present, since they do not accurately model multiple Coulomb scattering. To improve pencil beam dose calculations, methods for 2D pencil beam scaling were developed, as suggested by Szymanowski and Oelfke [99, 100]. These methods make use of the scattering power of a material, which can be modeled as a function of the CT numbers determined from an SECT scan. The scattering power of the material is used in these models to scale the dose from a pencil beam. This approach was shown to be more accurate in estimating the dose in the presence of lateral inhomogeneities. In

principle, the scattering power of a material can be determined from DECT, as done by Collins-Fekete *et al.* [101], and used as input for the pencil beam model. An alternative approach for pencil beam dose calculation was described by Yao *et al.* [102, 103], who uses a random walk model to describe the lateral dose from a pencil beam. Like the model proposed by Szymanowski and Oelfke, this approach makes use of the scattering power, which could be determined from DECT. It is left to future work to determine the impact of using DECT-determined scattering powers of materials for lateral scaling of a pencil beam.

To summarize, the work presented in this thesis provides a better understanding of three major sources of range uncertainties. Firstly, evidence for the benefits of DECT based tissue characterization for proton therapy planning is provided, and estimates of the related range uncertainties are given which can be used in clinical practice to estimate the safety margin. Secondly, an estimation of uncertainties from lateral inhomogeneities is provided. Thirdly, a rigorous estimation of the uncertainties related to the I-values is provided. The resulting values might be of interest for future clinical applications in MC dose calculations.

Bibliography

[1] Harald Paganetti. *Proton therapy physics*. CRC Press, 2016.

[2] American Society for Radiation Oncology (ASTRO). ASTRO Model Policies - Proton Beam Therapy, 2017.

[3] Particle Therapy Co-Operative Group (PTCOG). Patient statistics per end of 2016, 2016.

[4] Particle Therapy Co-Operative Group (PTCOG). Particle therapy facilitites in operation.

[5] E Rutherford. The scattering of α and β particles by matter and the structure of the atom. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 21(125):669–688, 1911.

[6] G Moliere. Theorie der Streuung schneller geladener Teilchen II Mehrfach-und Vielfachstreuung. *Zeitschrift fur Naturforschung - Section A Journal of Physical Sciences*, 1948.

[7] H Bethe. Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie. *Annalen der Physik*, 397(3):325–400, 1930.

[8] F Bloch. Zur Bremsung rasch bewegter Teilchen beim Durchgang durch Materie. *Annalen der Physik*, 1933.

[9] H Paganetti. Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Physics in medicine and biology*, 57(11):R99, 2012.

[10] U Schneider, E Pedroni, and A Lomax. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Physics in medicine and biology*, 41(1):111, 1996.

[11] W Schneider, T Bortfeld, and W Schlegel. Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions. *Physics in medicine and biology*, 45(2):459, 2000.

[12] E Pedroni, S Scheib, T Böhringer, A Coray, M Grossmann, S Lin, and A Lomax. Experimental characterization and physical modelling of the dose distribution of scanned proton pencil beams. *Physics in Medicine and Biology*, 50:541–561, 2005.

[13] A E Bourque, J-F Carrier, and H Bouchard. A stoichiometric calibration method for dual energy computed tomography. *Physics in medicine and biology*, 59(8):2059, 2014.

[14] D Han, J V Siebers, and J F Williamson. A linear, separable two-parameter model for dual energy CT imaging of proton stopping power computation. *Medical physics*, 43(1):600–612, 2016.

[15] N Hünemohr, B Krauss, C Tremmel, B Ackermann, O Jäkel, and S Greilich. Experimental verification of ion stopping power prediction from dual energy CT data in tissue surrogates. *Physics in medicine and biology*, 59(1):83, 2013.

[16] G Landry, J Seco, M Gaudreault, and F Verhaegen. Deriving effective atomic numbers from DECT based on a parameterization of the ratio of high and low linear attenuation coefficients. *Physics in medicine and biology*, 58(19):6851, 2013.

[17] C Möhler, P Wohlfahrt, C Richter, and S Greilich. Range prediction for tissue mixtures based on dual-energy CT. *Physics in medicine and biology*, 61(11):N268–N275, 2016.

[18] M Saito and S Sagara. A simple formulation for deriving effective atomic numbers via electron density calibration from dual-energy CT data in the human body. *Medical Physics*, 44(6):2293 –2303, 2017.

[19] V T Taasti, J B B Petersen, L P Muren, J Thygesen, and D C Hansen. A robust empirical parametrization of proton stopping power using dual energy CT. *Medical Physics*, 43(10):5547–5560, 2016.

[20] J K van Abbema, M-J van Goethem, M J W Greuter, A van der Schaaf, S Brandenburg, and E R van der Graaf. Relative electron density determination using a physics based parameterization of photon interactions in medical DECT. *Physics in medicine and biology*, 60(9):3825, 2015.

[21] M Saito. Potential of dual-energy subtraction for converting CT numbers to electron density based on a single linear relationship. *Medical physics*, 39(4):2021–2030, 2012.

[22] P Wohlfahrt, C Möhler, V Hietschold, S Menkel, S Greilich, M Krause, M Baumann, W Enghardt, and C Richter. Clinical implementation of dual-energy CT for proton treatment planning on pseudo-monoenergetic CT scans. *International Journal of Radiation Oncology* Biology* Physics*, 97(2):427–434, 2017.

[23] M Yang, G Virshup, J Clayton, X R Zhu, R Mohan, and L Dong. Theoretical variance analysis of single-and dual-energy computed tomography methods for calculating proton stopping power ratios of biological tissues. *Physics in medicine and biology*, 55(5):1343, 2010.

[24] N Hünemohr, H Paganetti, S Greilich, O Jäkel, and J Seco. Tissue decomposition from dual energy CT data for MC based dose calculation in particle therapy. *Medical physics*, 41(6):61714, 2014.

[25] G Landry, K Parodi, J E Wildberger, and F Verhaegen. Deriving concentrations of oxygen and carbon in human tissues using single-and dual-

energy CT for ion therapy applications. *Physics in medicine and biology*, 58(15):5029, 2013.

[26] A Lalonde and H Bouchard. A general method to derive tissue parameters for Monte Carlo dose calculation with multi-energy CT. *Physics in Medicine and Biology*, 61(22):8044, 2016.

[27] ICRU. Stopping Powers for Electrons and Positrons. *ICRU report 37*, 1984.

[28] Y Kumazaki, T Akagi, T Yanou, D Suga, Y Hishikawa, and T Teshima. Determination of the mean excitation energy of water from proton beam ranges. *Radiation Measurements*, 42(10):1683–1691, 2007.

[29] D Emfietzoglou, R Garcia-Molina, I Kyriakou, I Abril, and H Nikjoo. A dielectric response study of the electronic stopping power of liquid water for energetic protons and a new I-value for water. *Physics in medicine and biology*, 54(11):3451, 2009.

[30] T J Thompson. Effect of chemical structure on stopping powers for high-energy protons. 1952.

[31] J A Nordin and R M Henkelman. Measurement of stopping power ratios for 60 MeV positive or negative pions. *Physics in medicine and biology*, 24(4):781, 1979.

[32] H Bichsel and T Hiraoka. Energy loss of 70 MeV protons in elements. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 66(3):345–351, 1992.

[33] H Bichsel, T Hiraoka, and K Omata. Aspects of fast-ion dosimetry. *Radiation research*, 153(2):208–219, 2000.

[34] M J Berger and S M Seltzer. Stopping Powers and Ranges of Electrons and Positions. Technical report, 1982.

[35] M J Berger, M Inokuti, H H Andersen, H Bichsel, D Powers, S M Seltzer, D Thwaites, and D E Watt. Report 49. *Journal of the International Commission on Radiation Units and Measurements*, os25(2), may 1993.

[36] P Andreo. On the clinical spatial resolution achievable with protons and heavier charged particle radiotherapy beams. *Physics in medicine and biology*, 54(11):N205–N215, 2009.

[37] B Bednarz, J Daartz, and H Paganetti. Dosimetric accuracy of planning and delivering small proton therapy fields. *Physics in Medicine and Biology*, 55(24):7425–7438, 2010.

[38] G O Sawakuchi, U Titt, D Mirkovic, and R Mohan. Density heterogeneities and the influence of multiple Coulomb and nuclear scatterings on the Bragg peak distal edge of proton therapy beams. *Physics in medicine and biology*, 53(17):4605, 2008.

[39] J Schuemann, S Dowdell, C Grassberger, C H Min, and H Paganetti. Site-specific range uncertainties caused by dose calculation algorithms for proton therapy. *Physics in Medicine and Biology*, 59(15):4007–4031, 2014.

[40] H Paganetti, M Durante, C Zacharatou Jarlskog, C Lee, W E Bolch, B Bednarz, J Daartz, K Parodi, A Ferrari, F Sommerer, C Grassberger, A Lomax, and H Paganetti. Clinical implementation of full Monte Carlo dose calculation in proton beam therapy. *Phys. Med. Biol*, 53(17):4825 –4853, 2008.

[41] M Van Herk. Errors and Margins in Radiotherapy, 2004.

[42] J Leong. Implementation of random positioning error in computerised radiation treatment planning systems as a result of fractionation. *Physics in Medicine and Biology*, 1987.

[43] M Van Herk, P Remeijer, C Rasch, and J V Lebesque. The probability of correct target dosage: Dose-population histograms for deriving treatment margins in radiotherapy. *International Journal of Radiation Oncology Biology Physics*, 2000.

[44] ICRU. ICRU Report 50 - Prescribing, recording and reporting photon beam therapy. *ICRU report*, 1993.

[45] J C Stroom, H C J De Boer, H Huizenga, and A G Visser. Inclusion of geometrical uncertainties in radiotherapy treatment planning by means of coverage probability. *International Journal of Radiation Oncology Biology Physics*, 1999.

[46] E Bär, A Lalonde, R Zhang, K-W Jee, K Yang, G Sharp, B Liu, G Royle, H Bouchard, and H-M Lu. Experimental validation of two dual-energy CT methods for proton therapy using heterogeneous tissue samples. *Medical physics*, 45(1):48–59, 2018.

[47] A Lalonde, E Bär, and H Bouchard. A Bayesian approach to solve proton stopping powers from noisy multi-energy CT data. *Medical Physics*, 44(10):5293 –5302, 2017.

[48] D R White, H Q Woodard, and S M Hammond. Average soft-tissue and bone models for use in radiation dosimetry. *The British journal of radiology*, 60(717):907–913, 1987.

[49] H Q Woodard and D R White. The composition of body tissues. *The british journal of Radiology*, 59(708):1209–1218, 1986.

[50] B Schaffner and E Pedroni. The precision of proton range calculations in proton radiotherapy treatment planning: experimental verification of the relation between CT-HU and proton stopping power. *Physics in medicine and biology*, 43(6):1579, 1998.

[51] M Yang, X R Zhu, P C Park, U Titt, R Mohan, G Virshup, J E Clayton, and L Dong. Comprehensive analysis of proton range uncertainties related to patient stopping-power-ratio estimation using the stoichiometric calibration. *Physics in medicine and biology*, 57(13):4095, 2012.

[52] B J Heismann, J Leppert, and K Stierstorfer. Density and atomic number measurements with spectral x-ray attenuation method. *Journal of applied physics*, 94(3):2073–2079, 2003.

[53] M Bazalova, J-F Carrier, L Beaulieu, and F Verhaegen. Dual-energy CT-based material extraction for tissue segmentation in Monte Carlo dose calculations. *Physics in medicine and biology*, 53(9):2439–2456, 2008.

[54] J D Evans, B R Whiting, J A O'Sullivan, D G Politte, P H Klahr, Y Yu, and J F Williamson. Prospects for in vivo estimation of photon linear attenuation coefficients using postprocessing dual-energy CT imaging on a commercial scanner: Comparison of analytic and polyenergetic statistical reconstruction algorithms. *Medical physics*, 40(12):121914, 2013.

[55] J-É Tremblay, S Bedwani, and H Bouchard. A theoretical comparison of tissue parameter extraction methods for dual energy computed tomography. *Medical physics*, 41(8):81905, 2014.

[56] E C McCullough. Photon attenuation in computed tomography. *Medical Physics*, 2(6):307–320, 1975.

[57] R E Alvarez and A Macovski. Energy-selective reconstructions in x-ray computerised tomography. *Physics in medicine and biology*, 21(5):733, 1976.

[58] O Klein and Y Nishina. Über die Streuung von Strahlung durch freie Elektronen nach der neuen relativistischen Quantendynamik von Dirac. *Zeitschrift für Physik*, 52(11-12):853–868, 1929.

[59] R A Rutherford, B R Pullan, and I Isherwood. Measurement of effective atomic number and electron density using an EMI scanner. *Neuroradiology*, 11(1):15–21, 1976.

[60] D F Jackson and D J Hawkes. X-ray attenuation coefficients of elements and mixtures. *Physics Reports*, 70(3):169–233, 1981.

[61] M Torikoshi, T Tsunoo, M Sasaki, M Endo, Y Noda, Y Ohno, T Kohno, K Hyodo, K Uesugi, and N Yagi. Electron density measurement with dual-energy x-ray CT using synchrotron radiation. *Physics in medicine and biology*, 48(5):673, 2003.

[62] M J Berger Et al. XCOM: Photon Cross Section Database (version 1.5). Online, 2010.

[63] W V Mayneord. The significance of the roentgen. *Acta of the International Union against Cancer*, 2:271, 1937.

[64] ICRU. Tissue Substitutes in Radiation Dosimetry and Measurement. *ICRU report 44*, 1989.

[65] International Commission on Radiological Protection. Task Group and Walter Stephen Snyder. *Report of the task group on reference man*, volume 23. Pergamon Oxford, 1975.

[66] G Landry, F DeBlois, M Bazalova, and F Verhaegen. MO-D-303A-06: ImaSim, An Animated Tool for Teaching Imaging. *Medical Physics*, 36(6):2696–2697, 2009.

[67] G Landry, B Reniers, P V Granton, B van Rooijen, L Beaulieu, J E Wildberger, and F Verhaegen. Extracting atomic numbers and electron densities from a dual source dual energy CT scanner: experiments and a simulation model. *Radiotherapy and Oncology*, 100(3):375–379, 2011.

[68] B Schaffner. *Range precision of therapeutic proton beams.* PhD thesis, Diss. Naturwiss. ETH Zürich, Nr. 12474, 1997. Ref.: Peter Niederer; Korref.: Peter Rüegsegger; Korref.: Eros Pedroni, 1997.

[69] M J Berger. ESTAR, PSTAR, and ASTAR: Computer programs for calculating stopping-power and range tables for electrons, protons, and helium ions. *Unknown*, 1992.

[70] E Bär, A Lalonde, G Royle, H M Lu, and H Bouchard. The potential of dual-energy CT to reduce proton beam range uncertainties. *Med. Phys.*, 44(6):2332–2344, 2017.

[71] N Hudobivnik, F Schwarz, T Johnson, L Agolli, G Dedes, T Tessonnier, F Verhaegen, C Thieke, C Belka, and W H Sommer. Comparison of proton therapy treatment planning for head tumors with a pencil beam algorithm on dual and single energy CT images. *Medical physics*, 43(1):495–504, 2016.

[72] J Zhu and S N Penfold. Dosimetric comparison of stopping power calibration with dual energy CT and single energy CT in proton therapy treatment planning. *Medical physics*, 43(6):2845–2854, 2016.

[73] R Zhang, E Baer, K-W Jee, G C Sharp, J Flanz, and H-M Lu. Investigation of real tissue water equivalent path lengths using an efficient dose extinction method. *Physics in medicine and biology*, 62(14):5640–5651, 2017.

[74] M. Urie, M. Goitein, W. R. Holley, and G. T.Y. Chen. Degradation of the Bragg peak due to inhomogeneities. *Physics in Medicine and Biology*, 31(1):1–15, 1986.

[75] M J Berger. ESTAR, PSTAR and ASTAR: Computer programs for calculating stopping powers and ranges for electrons, protons and helium ions. 1995.

[76] G C Sharp, N Kandasamy, H Singh, and Michael Folkert. GPU-based streaming architectures for fast cone-beam CT image reconstruction and demons deformable registration. *Physics in medicine and biology*, 52(19):5771, 2007.

[77] ICRU. Stopping powers and ranges for protons and alpha particles. *ICRU report 49*, 1993.

[78] V De Smet, Labarbe R, F Vander Stappen, B Macq, and E Sterpin. A Water-Content Based Formalism to Mitigate Stopping Power Ratio Uncertainties Caused by Mean Excitation Energy Uncertainties in Hadron Therapy. *Medical Physics*, 2017.

[79] E Bär, P Andreo, A Lalonde, G Royle, and H Bouchard. Optimized I-values for use with the Bragg additivity rule and their impact on proton stopping power and range uncertainty. *Physics in Medicine and Biology*, 63(16):165007, 2018.

[80] G D Zeiss, William J Meath, J C F MacDonald, and D J Dawson. Accurate Evaluation of Stopping and Straggling Mean Excitation Energies for N, O, H2, N2, O2, NO, NH3, H2 O, and N2 O Using Dipole Oscillator Strength Distributions: A Test of the Validity of Bragg's Rule. *Radiation Research*, 70(2):284–303, 1977.

[81] B L Jhanwar, W J Meath, and J C F MacDonald. Dipole oscillator strength distributions and sums for C2H6, C3H8, n-C4H10, n-C5H12, n-C6H14, n-C7H16, and n-C8H18. *Canadian Journal of Physics*, 59(2):185–197, 1981.

[82] H Bichsel and R H Hilko. Measurement of the energy loss of alpha particles in carbon dioxide. *Helvetica Physica Acta*, 53(4):655, 1981.

[83] G F Thomas and W J Meath. Dipole spectrum, sums and properties of ground-state methane and their relation to the molar refractivity and dispersion energy constant. *Molecular Physics*, 34(1):113–125, 1977.

[84] M Bader, R E Pixley, F S Mozer, and W Whaling. Stopping cross section of solids for protons, 50-600 Kev. *Physical Review*, 103(1):32–38, 1956.

[85] L R Painter, E T Arakawa, M W Williams, and J C Ashley. Optical properties of polyethylene: Measurement and applications. *Radiation Research*, 83(1):1–18, 1980.

[86] C Tschalär and H Bichsel. Mean excitation potential of light compounds. *Physical Review*, 175(2):476, 1968.

[87] J C Ashley, C J Tung, and R H Ritchie. Inelastic interactions of electrons with polystyrene: calculations of mean free paths, stopping powers, and CSDA ranges. *IEEE Transactions on Nuclear Science*, 25(6):1566–1570, 1978.

[88] A Besemer, H Paganetti, and B Bednarz. The clinical impact of uncertainties in the mean excitation energy of human tissues during proton therapy. *Physics in medicine and biology*, 58(4):887–902, 2013.

[89] P J Doolan, C-A Collins-Fekete, M F Dias, Thomas A Ruggieri, D D'Souza, and J Seco. Inter-comparison of relative stopping power estimation models for proton therapy. *Physics in Medicine & Biology*, 61(22):8085, 2016.

[90] A Sudhyadhom. Determination of mean ionization potential using magnetic resonance imaging for the reduction of proton beam range uncertainties: theory and application. *Physics in Medicine & Biology*, 62(22):8521, 2017.

[91] T Hiraoka, K Kawashima, K Hoshino, A Fukumura, and Hans Bichsel. Energy loss of 70 MeV protons in organic polymers. *Medical physics*, 20(1):135–141, 1993.

[92] T Hiraoka, K Kawashima, K Hoshino, and Hans Bichsel. Energy loss of 70 mev protons in tissue-substitute materials. *Physics in medicine and biology*, 39(6):983, 1994.

[93] M J Berger, M Inokuti, H H Anderson, H Bichsel, J A Dennis, D Powers, S M Seltzer, and J E Turner. Report 37. *Journal of the International Commission on Radiation Units and Measurements*, os19(2), dec 1984.

[94] A Ivantchenko, V N Ivanchenko, J-M Quesada Molina, and S L Incerti. Geant4 hadronic physics for space radiation environment. *International Journal of Radiation Biology*, 88(1-2):171–175, 2012.

[95] ICRU. Key Data for Ionizing-Radiation Dosimetry: Measurement Standards and Applications. *ICRU report 90*, 2014.

[96] P Andreo, J Wulff, D T Burns, and H Palmans. Consistency in reference radiotherapy dosimetry: resolution of an apparent conundrum when 60Co is the reference quality for charged-particle and photon beams. *Physics in medicine and biology*, 58(19):6593–6621, 2013.

[97] M Goitein. Calculation of the uncertainty in the dose delivered during radiation therapy. *Medical Physics*, 12(5):608–612, 1985.

[98] H Palmans and F Verhaegen. Assigning nonelastic nuclear interaction cross sections to Hounsfield units for Monte Carlo treatment planning of proton beams. *Physics in Medicine and Biology*, 43:2755 –2768, 2005.

[99] H Szymanowski and U Oelfke. CT calibration for two-dimensional scaling of proton pencil beams. *Physics in Medicine and Biology*, 48:861–874, 2003.

[100] H Szymanowski and U Oelfke. Two-dimensional pencil beam scaling: an improved proton dose algorithm for heterogeneous media. *Physics in Medicine and Biology*, 47(18):3313–3330, 2002.

[101] C A Collins-Fekete, E Bär, L Volz, H Bouchard, L Beaulieu, and J Seco. Extension of the Fermi-Eyges most-likely path in heterogeneous medium with prior knowledge information. *Physics in Medicine and Biology*, 62:9207 –9219, 2017.

[102] W Yao, T E Merchant, and J B Farr. A simplified analytical random walk model for proton dose calculation. *Physics in Medicine and Biology*, 61:7412–7426, 2016.

[103] W Yao, T E Merchant, and J B Farr. A correction scheme for a simplified analytical random walk model algorithm of proton dose calculation in distal Bragg peak regions. *Physics in Medicine and Biology*, 61:7397–7411, 2016.